Asymptotics for the standard block size in primal lattice attacks: second order, formally verified

Daniel J. Bernstein

University of Illinois at Chicago, Department of Computer Science, USA
Ruhr University Bochum, Horst Görtz Institute for IT Security, Germany
Academia Sinica, Institute of Information Science, Taiwan

Abstract. Many proposals of lattice-based cryptosystems estimate security levels by following a recipe introduced in the New Hope proposal. This recipe, given a lattice dimension n, modulus q, and standard deviation s, outputs a "primal block size" β and a security level growing linearly with β. This β is minimal such that some κ satisfies $((n+\kappa)s^2+1)^{1/2} < (d/\beta)^{1/2}\delta^{2\beta-d-1}q^{\kappa/d}$, where $d = n+\kappa+1$ and $\delta = (\beta(\pi\beta)^{1/\beta}/(2\pi\exp 1))^{1/2(\beta-1)}$.

This paper identifies how β grows with n, with enough precision to show the impact of adjusting q and s by constant factors. Specifically, this paper shows that if $\lg q$ grows as $Q_0 \lg n + Q_1 + o(1)$ and $\lg s$ grows as $S_0 \lg n + S_1 + o(1)$, where $0 \le S_0 \le 1/2 < Q_0 - S_0$, then β/n grows as $z_0 + (z_1 + o(1))/\lg n$, where $z_0 = 2Q_0/(Q_0 - S_0 + 1/2)^2$ and z_1 has a formula given in the paper. The paper provides a traditional-format proof and a proof verified by the HOL Light proof assistant.

1 Introduction

The order of growth of the running time of an algorithm, defined in Chapter 2, gives a simple characterization of the algorithm's efficiency and also allows us to compare the relative performance of alternative algorithms. Once the input size n becomes large enough, merge sort, with its $\Theta(n \lg n)$ worst-case running time, beats insertion sort, whose worst-case running time is $\Theta(n^2)$. Although we can sometimes determine the exact running time of an algorithm, as we did for insertion sort in Chapter 2, the extra precision is not usually worth the effort of computing it. For large enough inputs, the multiplicative constants and lower-order terms of an exact running time are dominated by the effects of the input size itself.
—Cormen–Leiserson–Rivest–Stein, "Introduction to algorithms" [**C9LRS09**, page 43]

Assessing algorithmic performance makes use of the "Big Oh" notation that proves essential to compare algorithms, and design better ones.
—Skiena, "The algorithm design manual" [**S2.20**, page 31]

This work was funded by the Intel Crypto Frontiers Research Center; by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of the Excellence Strategy of the German Federal and State Governments—EXC 2092 CASA—390781972 "Cyber Security in the Age of Large-Scale Adversaries"; by the U.S. National Science Foundation under grant 1913167; and by the Taiwan's Executive Yuan Data Safety and Talent Cultivation Project (AS-KPQ-109-DSTCP). "Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation" (or other funding agencies). Permanent ID of this document: `22882984e77307356f7598b7ac1c434d8f31613f`. Date: 2024.04.27.

E-mail: djb@cr.yp.to (Daniel J. Bernstein)

This is a paper on the analysis of attacks against post-quantum cryptography, but the goals of the paper are most easily understood by analogy to various known results on pre-quantum cryptography, so the paper begins with those results. Beware that many of the following statements are conjectures relying on heuristics.

The number-field sieve (NFS) factors an integer n in time subexponential in the number of bits of n, more precisely $\exp((\log n)^{1/3+o(1)})$ bit operations. This formula is communicating important information about scalability. For example:

- RSA needs $\lambda^{3+o(1)}$ bits in n to reach λ bits of security against NFS. For comparison, elliptic-curve cryptography (ECC) using appropriately chosen curves over $\mathbb{F}_p$ reaches λ bits of security against known attacks with p having just $\lambda^{1+o(1)}$ bits, and with user computations taking just $\lambda^{2+o(1)}$ bit operations, evidently smaller and faster than RSA for all sufficiently large λ.

- Before NFS, many factorization algorithms were known that used $\exp((\log n)^{1/2+o(1)})$ bit operations—in some cases provably so; see the survey in [**L5P92**, Section 1]. Initial analyses of NFS indicated that it would be slower than earlier algorithms for all sizes of n of interest (see, e.g., [**A2.91**, page 69, "no practical value when the number to be factored is not of special form"]), but the fact that NFS had improved 1/2 to 1/3 continued to drive research into the details of NFS (see, e.g., [**L4L93**]), and eventually NFS was setting factorization records (see, e.g., [**C4DLL+00**]).

- Coppersmith [**C8.84**] had achieved cost $\exp((\log n)^{1/3+o(1)})$ earlier for a related problem, namely computing discrete logarithms in $\mathbb{F}_n^*$ when n is a power of 2. Many years later, Joux in [**J14**] achieved cost $\exp((\log n)^{1/4+o(1)})$ for that problem, and the change from 1/3 to 1/4 turned out to be the first step towards a quasi-polynomial-time algorithm for that problem. See generally [**B1GJE14**] and [**G2KZ18**].

But most work in cryptography relies on more detailed analyses of attack cost. As an example of why more detailed analyses are needed, imagine someone taking the statement that ECC reaches λ bits of security with $\lambda^{1+o(1)}$ bits in p; incorrectly substituting $o(1) = 0$; and then taking $p \approx 2^{128}$ for 128 bits of security. Large-scale attackers can easily break this size of p using known "square-root" attacks; see [**OW99**]. Even worse, breaking 2^{40} targets costs only 2^{20} times as much as breaking one target; see [**K7S01**].

Asymptotically, the square-root attacks take $p^{1/2+o(1)}$ bit operations (as the name would suggest), limiting the security level to $(1/2 + o(1)) \lg p$, where as usual $\lg p$ means $\log_2 p$. For this security level to reach λ, one needs $\lg p$ to reach $(2 + o(1))\lambda$. Keys then have $(4 + o(1))\lambda$ bits if a curve point is sent as two coordinates, and have $(2 + o(1))\lambda$ bits in compressed form.

A similarly precise statement about RSA key sizes involves longer formulas. Specifically, NFS uses $\exp((c + o(1))(\log n)^{1/3}(\log\log n)^{2/3})$ bit operations for a particular constant c. A short calculation then says that RSA needs $\lg n$ to grow as $\lambda^3/((9c^3 + o(1))(\lg \lambda)^2)$.

Lenstra–Lenstra–Manasse–Pollard [**L4LMP93**, Section 9] outlined a preliminary version of NFS with $c^3 = 9 \approx 2.080^3$. Buhler–Lenstra–Pomerance [**B6LP93**] presented a better version of NFS (using an idea from [**A2.91**]) improving c^3 to $64/9 \approx 7.111 \approx 1.923^3$. Coppersmith [**C8.93**] further improved c^3 to $(92 + 26\sqrt{13})/27 \approx 6.879 \approx 1.902^3$, and to $(20 + 8\sqrt{6})/9 \approx 4.400 \approx 1.639^3$ in situations where one has a large batch of RSA keys to factor at once. This level of precision also allows analysis of the impact of accounting for, and of optimizing, memory-access costs inside attacks on a two-dimensional or three-dimensional circuit; see, e.g., [**B4L14**].

Much more of the literature becomes visible when one carries out even more precise analyses. For example, $0.5 + o(1)$ bits of security were shaved off ECC by the "negation" speedup; see [**B4LS11**] and the references therein. Seeing this speedup requires paying attention to constant factors in attack cost, unlike the above quote from [**C9LRS09**].

For RSA, there are many years of papers such as [C4DLL+00], [K4AFL+10], and [B5GGH+20] building complete software for more and more advanced versions of NFS, and relying on a computer to monitor the cost of running each algorithm. Such experimental monitoring can convincingly produce precise, accurate average-case observations—certainly much better than incorrectly substituting $o(1) = 0$ into formulas such as $\exp((c + o(1))(\log n)^{1/3}(\log\log n)^{2/3})$. Similar comments apply to various other attack algorithms.

Experiments are not a substitute for asymptotic analysis; they are complements to asymptotic analysis. As NFS illustrates, the algorithms that perform best in experiments—or that are convincingly argued to be the best options for large-scale attacks and thus to dictate security levels even when they are not the best in small-scale experiments—often arise from a multi-stage process of (1) pursuing better asymptotics and then (2) working on lower-order speedups to the asymptotically best algorithms.

1.1 Asymptotics for lattice attacks

Moving from RSA and ECC to lattice-based cryptography makes it much more difficult to find clear statements of how large key sizes need to be to achieve λ bits of security against known attacks. Some *components* of the necessary analysis are available and are summarized in the following paragraphs; as before, many of these statements are conjectures relying on heuristics.

One component is asymptotic analysis of the cost of solving n-dimensional shortest-vector problems. A lattice-sieving algorithm from [N1V08] uses $2^{(\alpha+o(1))n}$ bit operations (assuming the input lattice is specified by a matrix of integers having $2^{o(n)}$ bits) for a particular constant α, namely $\lg(4/3) \approx 0.415$. A series of followup papers [WLTB11], [ZPH14], [B2GJ14], [L1.15], [L1W15], and [B2DGL16] improved α to $\lg\sqrt{3/2} \approx 0.292$, or $\lg\sqrt{13/9} \approx 0.265$ for quantum attacks. It was reported in [K3L21] that 0.292 and 0.265 were optimal "within a broad class of lattice sieving algorithms covering almost all approaches to date"; [C5L21] and [H3.21] both reported quantum attacks with α below 0.265, but there have been no improvements to the 0.292.

It is important to realize that typical cryptosystems using n-dimensional lattices pose attack problems that are not the n-dimensional shortest-vector problem. The security levels are only a fraction of $(\alpha + o(1))n$; the fraction depends on cryptosystem parameters.

As a specific target for attacks, Section 2.1 reviews the Lyubashevsky–Peikert–Regev PKE [L7PR13, eprint version, page 4], using the ring $(\mathbb{Z}/q)[x]/(x^n + 1)$ and sampling each error position from distribution χ. The standard heuristic analysis cares only about (n, q, s), where s is the standard deviation of χ.

Section 2.2 reviews the standard "primal" key-recovery attack. A critical parameter in this attack is a "block size" β. The attack takes time $2^{(\alpha+o(1))\beta}$ for a BKZ-β computation, which internally solves $2^{o(\beta)}$ shortest-vector problems, each β-dimensional.

There is a standard recipe to select β given (n, q, s) as input: for each integer $\beta \geq 60$ in turn, try each $\kappa \in \{1, 2, \ldots, n\}$, and stop as soon as $\text{StandardRatio}(n, q, s, \kappa, \beta) < 1$, where StandardRatio is defined as follows.

Definition 1.1.1 (the standard ratio). *Let n, q, s, κ, β be real numbers such that $2 \leq n$; $2 \leq q$; $0 < s$; $1 \leq \kappa$; and $2 \leq \beta$. Then* $\text{StandardRatio}(n, q, s, \kappa, \beta)$ *is defined as $((n+\kappa)s^2 + 1)^{1/2}/(d/\beta)^{1/2}\delta^{2\beta-d-1}q^{\kappa/d}$ where $d = n + \kappa + 1$ and $\delta = (\beta(\pi\beta)^{1/\beta}/(2\pi\exp 1))^{1/2(\beta-1)}$.*

This recipe was introduced by Alkim–Ducas–Pöppelmann–Schwabe [A4DPS16, Section 6.3] as part of the New Hope proposal. The analysis from [A4DPS16] says that the standard attack works using attack parameters κ and β if and only if $\text{StandardRatio}(n, q, s, \kappa, \beta) < 1$; see Section 2.3, and see Section 2.4 regarding the 60.

There have been subsequent speedups and corrections, but the literature generally portrays these adjustments as minor; see, e.g., [A3D21], starting with the "history of refinements" title. The same recipe has been applied by many newer cryptosystem proposals; see, e.g., [A4BDL+21, Section 5.2.2], [ETWY22, Section 4.3], and [C6CHY24, Table 2].

Note that this recipe outputs only *examples* of β, not the *asymptotics* of β. An asymptotic formula would say how β grows with n, where the coefficients in the growth rate are functions of variables specifying how q and s grow with n. The next paragraph gives three examples of questions that one would expect to easily answer by looking at an asymptotic formula, and that do not seem to be answered in the literature.

As context for the questions, designers of lattice-based encryption systems (outside the context of homomorphic encryption) typically take q somewhere between $n^{1/2+o(1)}$ and $n^{1+o(1)}$ and take s to be constant. Decryption relies on q being large by comparison to s, but cryptosystem details (for example, the amount of "error correction") influence how large, and influence the choice of s. Here are the questions: If a cryptosystem modification replaces s with $2s$, what is the asymptotic impact on β? What about replacing q with $q/2$? What about moving q from $n^{1+o(1)}$ all the way down to $n^{1/2+o(1)}$?

1.2 Contributions of this paper

The point of this paper is to calculate how the standard choice of β scales with the cryptosystem parameters (n, q, s), with enough detail to see the asymptotic impact of multiplying q or s by a constant factor. The main theorem is as follows.

Theorem 1.2.1 (asymptotic growth of the standard block size). *Let Q_0, Q_1, S_0, S_1 be real numbers such that $0 \le S_0 \le 1/2 < Q_0 - S_0$. Let N be an infinite subset of $\{2, 3, 4, 5, \ldots\}$. Let $n \mapsto q$ and $n \mapsto s$ be functions from N to $\mathbb{R}$ such that*

$$\begin{aligned} 2 \le q, &\qquad \lg q \in Q_0 \lg n + Q_1 + o(1), \\ 0 < s, &\qquad \lg s \in S_0 \lg n + S_1 + o(1). \end{aligned}$$

Define $x_0 = (Q_0 + S_0 - 1/2)/(Q_0 - S_0 + 1/2)$; $z_0 = 2Q_0/(Q_0 - S_0 + 1/2)^2$; and

$$z_1 = \left(2S_1 + \lg z_0 - \left(S_0 - Q_0 + \frac{3}{2}\right) \lg \frac{z_0}{2\pi \exp 1} - \frac{Q_1(Q_0 + S_0 - \frac{1}{2})}{Q_0}\right) \frac{2Q_0}{(Q_0 - S_0 + \frac{1}{2})^3}.$$

(1) There are functions $n \mapsto \kappa$ and $n \mapsto \beta$ from N to $\mathbb{Z}$ such that

$$\begin{aligned} &1 \le \kappa \le n \text{ for all } n, && \kappa/n \in x_0 + o(1)/\lg n, \\ &2 \le \beta \le n + \kappa + 1 \text{ for all } n, && \beta/n \in z_0 + (z_1 + o(1))/\lg n, \qquad \text{and} \\ &\text{StandardRatio}(n, q, s, \kappa, \beta) < 1 \text{ for all sufficiently large } n. \end{aligned}$$

(2) Let $n \mapsto \kappa$ and $n \mapsto \beta$ be functions from N to $\mathbb{R}$ such that

$$\begin{aligned} &1 \le \kappa \le 100n \text{ for all } n, \\ &60 \le \beta \le n + \kappa + 1 \text{ for all } n, \qquad \text{and} \\ &\text{StandardRatio}(n, q, s, \kappa, \beta) \le 1 \text{ for all sufficiently large } n. \end{aligned}$$

Then $\beta \ge \ell$ for some function $n \mapsto \ell$ with $\ell/n \in z_0 + (z_1 + o(1))/\lg n$.

In short, the standard block size β has β/n growing as $z_0 + (z_1 + o(1))/\lg n$, where z_0 and z_1 are given by the formulas in the theorem statement.

This paper includes two proofs of the theorem: a proof in a traditional format, and a proof verified by the HOL Light [H2.96] proof assistant. Section 3 presents the

traditional-format proof. Appendix A explains the value of formal verification in this context, compares the formally verified theorem statement line by line to the statement of Theorem 1.2.1, and explains how to run HOL Light to re-check the formally verified proof. The lemmas and proof steps in Section 3 should be viewed as having lower assurance, since they have not been compared line by line to formally verified statements.

The reader is cautioned that the standard block size could be different from, perhaps far from, the block size that is actually required by the standard primal attack. This paper is rigorously analyzing the asymptotics of the standard block size; as context, this paper reviews the heuristic analysis from [**A4DPS16**] saying that the standard block size is the required block size; this paper should not be interpreted as endorsing, or providing any evidence for, the analysis from [**A4DPS16**].

1.3 Examples of using the asymptotics

One immediately sees from Theorem 1.2.1 that z_0 is independent of S_1, and that z_1 is $4S_1Q_0/(Q_0 - S_0 + 1/2)^3$ plus something independent of S_1. This answers the question of what happens if s is replaced by $2s$: the change increases S_1 by 1, so it increases z_1 by $4Q_0/(Q_0-S_0+1/2)^3$, so it increases β/n by $(4Q_0/(Q_0-S_0+1/2)^3+o(1))/\lg n$. For example, the increase in β/n is $(32/27 + o(1))/\lg n$ for $(Q_0, S_0) = (1, 0)$, or $(192/125 + o(1))/\lg n$ for $(Q_0, S_0) = (3/4, 0)$.

Similar comments apply to the question of what happens if q is replaced with $q/2$. This decreases Q_1 by 1, so it increases z_1 by $2(Q_0 + S_0 - 1/2)/(Q_0 - S_0 + 1/2)^3$, so it increases β/n by $(2(Q_0 + S_0 - 1/2)/(Q_0 - S_0 + 1/2)^3 + o(1))/\lg n$. For example, the increase in β/n is $(8/27 + o(1))/\lg n$ for $(Q_0, S_0) = (1, 0)$, or $(32/125 + o(1))/\lg n$ for $(Q_0, S_0) = (3/4, 0)$.

There is an indirect effect of replacing q with $q/2$ if key sizes are held constant. The A component of an LPR public key (see Section 2.1) has $n \lg q \in n(Q_0 \lg n + Q_1 + o(1))$ bits, so decreasing Q_1 by 1 saves $(1 + o(1))n$ bits.[1] This in turn allows n to be increased to $n+n/((Q_0+o(1)) \lg n)$ for the same key size, also multiplying β by $1+1/((Q_0+o(1)) \lg n)$, i.e., increasing β/n by $(z_0/Q_0 + o(1))/\lg n = (2/(Q_0 - S_0 + 1/2)^2 + o(1))/\lg n$. This indirect increase in β/n is $(8/9+o(1))/\lg n$ for $(Q_0, S_0) = (1, 0)$, or $(32/25+o(1))/\lg n$ for $(Q_0, S_0) = (3/4, 0)$. Note that, at this level of detail, the total of the direct and indirect effects of replacing q with $q/2$ matches the effect of doubling s.

As a simpler example, write a for the number of bits in A, and consider the question of how a scales to first order with the security level λ. Here it suffices to use $\beta \in (z_0 + o(1))n$, so $\lambda \in (\alpha z_0 + o(1))n$; this also implies $\lg \lambda \in (1 + o(1)) \lg n$, so $\lambda \lg \lambda \in (\alpha z_0 + o(1))n \lg n$. Meanwhile $a \in (Q_0 + o(1))n \lg n$, so $a/(\lambda \lg \lambda) \in Q_0/\alpha z_0 + o(1)$; i.e.,

$$a \in \left(\frac{(Q_0 - S_0 + 1/2)^2}{2\alpha} + o(1) \right) \lambda \lg \lambda.$$

For example, if $Q_0 - S_0 = 1$, then A has $(9/8\alpha+o(1))\lambda \lg \lambda$ bits: e.g., $(3.846\ldots+o(1))\lambda \lg \lambda$ bits for $\alpha = \lg \sqrt{3/2} \approx 0.292$. If a cryptosystem reduces $Q_0 - S_0$ to $1/2 + \epsilon$ then the number of bits in A drops to $((1 + \epsilon)^2/2\alpha + o(1))\lambda \lg \lambda$, reducing asymptotic key sizes by a factor close to $9/4$.

As a different application of the second-order formulas, consider the following question: if two cryptosystems both have $Q_0 - S_0 = 1$ and $Q_1 = S_1$ (so $\lg q - \lg s \in \lg n + o(1)$), but the first cryptosystem has $S_0 = 0$ while the second has $S_0 = 1/2$, then which cryptosystem has an asymptotically smaller ratio $a/(\beta \lg \beta)$, where a is the number of bits in A?

To answer this question, first use $\beta \in n(z_0+(z_1+o(1))/\lg n)$ to obtain $\lg \beta \in \lg n+\lg z_0+o(1)$ and $\beta \lg \beta \in (n \lg n)(z_0 + (z_1 + z_0 \lg z_0 + o(1))/\lg n)$. Also $a \in n(Q_0 \lg n + Q_1 + o(1))$, so the target ratio $a/(\beta \lg \beta)$ is $(Q_0 + (Q_1 + o(1))/\lg n)/(z_0 + (z_1 + z_0 \lg z_0 + o(1))/\lg n)$.

[1] An LPR public key also has a component G, but typically G is either shared across keys or generated from a $\Theta(\lambda)$-bit seed in each key; either way, the effect of q on the size of G is not relevant to key size.

Substituting $Q_0 = S_0 + 1$ and $Q_1 = S_1$ gives, after a short calculation, target ratio $9/8 - ((3/2)\lg z_0 + (3/8)\lg(2\pi \exp 1) + o(1))/\lg n$, which is asymptotically smaller when $z_0 = (8/9)Q_0$ is larger. In particular, the first cryptosystem has ratio $9/8 - (1.280\ldots + o(1))/\lg n$ while the second cryptosystem has asymptotically smaller ratio $9/8 - (2.157\ldots + o(1))/\lg n$.

This last example should not be taken as suggesting that the common practice of taking s to be small is suboptimal. Increasing s, while preserving q/s, tends to make legitimate decryption more difficult; i.e., the second type of cryptosystem tends to be harder to build than the first.

1.4 Related work

A search did not find any previous literature giving simple first-order asymptotic descriptions of lattice key sizes and the effect of error correction, never mind the second-order calculations that are the main work in this paper. The closest related work appears to be [**A3.17**, bottom of page 15], which, in the case $S_0 = 0$ and $Q_0 \in \mathbb{Z}$, carries out a first-order comparison (in a different model; see below) of two lattice attacks, concluding that the exponent for one attack is asymptotically $Q_0/(Q_0 + 1/2)$ times the exponent for the other, without stating an asymptotic formula for the performance of either attack.

The literature does present asymptotics for sieving cost in terms of β, but only the first-order asymptotics $2^{(\alpha+o(1))\beta}$ mentioned above. Literature presenting second-order improvements in β (for example, "dimensions for free" from [**D3.18**] reduce β to $\beta - (\lg(4/3) + o(1))\beta/\lg\beta$; the improved techniques from [**D2LW20**] improve $\lg(4/3) \approx 0.415$ to $\lg(13/9) \approx 0.531$) does not give similarly precise formulas for how β depends on cryptosystem parameters before or after the improvement, or for the resulting attack cost. A second-order analysis saying that sieving costs $2^{(\alpha_0+(\alpha_1+o(1))/\lg\beta)\beta}$ would compose with this paper's second-order analysis of how β depends on (n, q, s), and would allow second-order analysis of how key sizes scale with λ.

The standard block size from [**A4DPS16**] is not the only model in the literature of the block size that is actually required by the standard attack. For example, the analysis in [**A3.17**] used an earlier model of the required block size. As another example, some cryptosystem proposals, such as [**B4CLV18**], rely on "simulators" that (1) have more convincing justifications than the standard model and (2) are experimentally observed to produce more accurate results than the standard model for small-scale attacks. This paper's computation of the second-order asymptotics of the standard block size is a step towards comparing asymptotics of multiple models.

Beyond the primal attacks covered in this paper, it would be interesting to analyze the asymptotics of "dual" attacks. The literature on dual attacks is unsettled at the moment. Older analyses indicated that dual attacks (analyzed in a similar way to primal attacks) could sometimes outperform primal attacks; this was challenged in [**A5BDK+20**, page 26, "Primal attack only"]; newer dual attacks were reported in [**G3J21**] and [**M3.22**]; those attacks were challenged in [**D3P23**]; [**P1S23**] indicates that a modified attack can work around the issue raised in [**D3P23**]; it is unclear how that attack compares to primal attacks.

1.5 Priority dates

This paper's main theorem and formally verified proof were originally posted in March 2023 as part of a larger paper, "Multi-ciphertext security degradation for lattices" [**B4.23**]. Some components of the calculations were already posted in November 2022 as part of the first version of that paper.

I have decided to split this material out of that paper; that paper will be revised correspondingly to cite this paper for this material.

2 The standard block size

This section reviews the motivation in the literature for studying the standard block size. In particular, this section explains the components of Definition 1.1.1.

As noted in Section 1.1, this paper takes the LPR cryptosystem as a concrete target. Section 2.1 reviews this PKE. Section 2.2 reviews the standard "primal" key-recovery attack against this PKE. Section 2.3 reviews the standard analysis of the attack. Section 2.4 reviews a known flaw in the analysis for small values of β. Section 2.5 reviews the standard primal message-recovery attack. Section 2.6 reviews various other attacks, including attacks not covered by the standard analysis.

As illustrated by the tables in [**A3CDD+18**], the standard analysis has also been applied to a wide range of further cryptosystems. All the analysis needs to know about each cryptosystem is the lattice dimension n, the modulus q, and the standard deviation s. Complications such as the error-correcting codes in New Hope or the matrices in Kyber interact with the analysis only in how they end up choosing (n, q, s).

2.1 Review of the LPR cryptosystem

This PKE has three parameters: an integer $n \geq 2$; an integer $q \geq 2$; and a probability distribution χ supported on a finite set of integers. Assume for simplicity that the average of χ is 0. Write R for the ring $\mathbb{Z}[x]/(x^n + 1)$.

Key generation works as follows. Generate uniform random $G \in R/q$. Generate $a, e \in R$ with coefficients drawn independently at random from χ. Compute $A = aG + e \in R/q$. The secret key is (a, e). The public key is $(G, A) \in (R/q)^2$.

The set of messages is the set of elements of R with coefficients in $\{0, \lceil q/2 \rceil\}$. Encryption of a message M to a public key (G, A) works as follows. Generate $b, c, d \in R$ with coefficients drawn independently at random from χ. Compute $B = Gb + d \in R/q$ and $C = M + Ab + c \in R/q$. The ciphertext is $(B, C) \in (R/q)^2$.

Decryption of a ciphertext (B, C) works as follows. Compute $X = C - aB \in R/q$. Round each coefficient of X to the closest element of $\{0, \lceil q/2 \rceil\}$ in $\mathbb{Z}/q$, specifically 0 if both elements are equally close.[2]

The above PKE definition skips two requirements from the LPR paper, namely that n is a power of 2 and that q is a prime congruent to 1 modulo $2n$; see [**L7PR13**, Section 1.1]. Cryptosystems after [**L7PR13**] loosened the restrictions on q; for example, Kyber's current prime 3329 is $1 + 13 \cdot 256$. As for n, readers concerned about attacks enabled by factors of $x^n + 1$ in $\mathbb{Z}[x]$ should feel free to substitute the marginally larger polynomial $x^n - x - 1$, as in [**B4CLV18**]; this is orthogonal to the topic of this paper.

Correct decryption requires $X = C - aB = M + eb + c - ad$ to round to M, i.e., requires each coefficient of $eb + c - ad$ to be smaller than about $q/4$. If χ is, e.g., the uniform distribution on $\{-1, 0, 1\}$ then each coefficient of eb is a sum of n products where one expects about 4/9 to be nonzero, evenly balanced between 1 and -1, so typically the coefficient will be on the scale of $\sqrt{n}$, with considerable variation in the exact size. There are various proposals to reduce q close to this scale, and to avoid frequent decryption failures by applying an error-correcting code to M. In the opposite direction, sometimes cryptosystems take larger χ and correspondingly larger q; sometimes cryptosystems pack more message bits into each coefficient, again taking larger q. To cover many different cases, this paper considers a spectrum of possibilities for the asymptotic sizes of q and s.

[2] This rounding detail is not specified in [**L7PR13**]; also, [**L7PR13**] says $\lfloor q/2 \rceil$ without specifying whether the rounding rounds 0.5 up or rounds to even. These details do not affect the standard analysis; they are specified here so as to have a complete PKE definition.

2.2 Review of the standard key-recovery attack

Consider the problem of recovering the private key $(a, e) \in R^2$ from the public key $(G, A) \in (R/q)^2$. Recovering a suffices, since $e = A - aG$ by definition. The standard "primal" attack works as follows.

There is an attack parameter $\kappa \leq n$. Define a function $\text{First}_\kappa : R \to \mathbb{Z}^\kappa$ that extracts the first κ coefficients from its input. This induces a function, also written First_κ, from R/q to $(\mathbb{Z}/q)^\kappa$.

Define L as the set of all $(\alpha, \epsilon, \gamma) \in R \times \mathbb{Z}^\kappa \times \mathbb{Z}$ such that $\text{First}_\kappa(\gamma A - \alpha G)$ is the same as ϵ modulo q. This is a lattice of full rank $d = n + \kappa + 1$ and determinant q^κ. Note that $\pm(a, \text{First}_\kappa(e), 1)$ are elements of this lattice.

There is another attack parameter β. The attack writes down a basis for L, applies BKZ-β to reduce this basis, and hopes that BKZ-β outputs one of the short nonzero vectors $\pm(a, \text{First}_\kappa(e), 1)$, in particular revealing a.

The problem being attacked here, the problem of finding a, e given a random G and $aG + e$, is typically called "Ring-LWE", specifically "normal-form 1-sample search Ring-LWE", where "normal form" refers to the secret a being small. The Ring-LWE problem is typically credited to [**S4STX09**] and [**L7PR13**]. However, this problem was already attacked in the 1998 Hoffstein–Pipher–Silverman NTRU paper, both for the homogeneous case $A = 0$ (see [**H4PS1998**, Section 3.4.1]) and for general A (see [**H4PS1998**, Section 3.4.2]). The problem statements in [**S4STX09**] and [**L7PR13**] merely generalize to more "samples": e.g., finding a, e_1, e_2 given random $G_1, G_2, aG_1 + e_1, aG_2 + e_2$, or equivalently replacing $G \in R/q$ and $e \in R$ with row vectors $(G_1\ G_2) \in (R/q)^2$ and $(e_1\ e_2) \in R^2$ respectively.

The attack in [**H4PS1998**] has $\kappa = n$. May–Silverman [**M4S01**] generalized the attack to any $\kappa \leq n$. In the original 1996 NTRU handout [**H4PS2016**], various concrete examples chose different sizes for a and e, motivating another generalization from Coppersmith and Shamir [**C8S97**] to set up a lattice with, e.g., short vector $(3.14a, e, 1)$ rather than $(a, e, 1)$; this paper focuses on cryptosystems that take a and e of the same size, such as the LPR system. There can still be a tiny improvement from setting up a lattice with, e.g., short vector $(a, e, 3.14)$; this paper ignores this improvement for simplicity.

2.3 Review of the standard analysis

Once β is reasonably large, the main bottleneck in the standard attack is the BKZ-β computation. Part of the standard analysis is an analysis of the cost of BKZ-β. One complication here is that BKZ-β is a family of algorithms, not a single algorithm. One reason for improvements in the cost of BKZ-β is that underlying subroutines have been improved, notably for SVP-β; see Section 1.1. Another reason is that the use of those subroutines inside BKZ-β has been improved.

The rest of the standard analysis focuses on the question of how large β needs to be for BKZ-β to succeed at finding the target vector. The standard heuristic conclusion is that BKZ-β succeeds if and only if the 2-norm of the target vector is below $(d/\beta)^{1/2}\delta^{2\beta-d-1}q^{\kappa/d}$, where $\delta = (\beta(\pi\beta)^{1/\beta}/(2\pi \exp 1))^{1/2(\beta-1)}$. The rationale for this inequality is as follows:

- One heuristic says that, for a "random" lattice L of rank d, BKZ-β finds a nonzero vector of length close to $\delta^{d-1}(\det L)^{1/d}$, with δ defined as above.
- Another heuristic says that the Gram–Schmidt lengths of the BKZ-β output are close to a geometric series. Combining this with the shortest length being close to $\delta^{d-1}(\det L)^{1/d}$ and the product of the lengths being $\det L$ says approximately how large each length is. In particular, the length at position $d - \beta + 1$ is close to $\delta^{2\beta-d-1}(\det L)^{1/d}$, which for this lattice is $\delta^{2\beta-d-1}q^{\kappa/d}$. The rationale treats this approximation as an equation.

- Another heuristic says that if the target vector has length t then its projection onto the space spanned by the last β Gram–Schmidt vectors has length approximately $t\sqrt{\beta/d}$. The rationale also treats this approximation as an equation.
- If the latter length $t\sqrt{\beta/d}$ is below the previous length $\delta^{2\beta-d-1}q^{\kappa/d}$ then the above heuristics seem to contradict each other, since the last SVP-β call in each "tour" of BKZ-β guarantees that the projection of the vector at position $d-\beta+1$ is a minimum-length nonzero vector in the projection of L. Note, however, that the first heuristic was only for a "random" lattice. Another heuristic says that this seeming contradiction occurs if and only if BKZ-β detects the non-"randomness" of the lattice by finding the projection of v.
- A further heuristic says that BKZ-β finds the projection of v if and only if BKZ-β finds v. A slightly different statement appears in [**A3GVW17**], which says that if BKZ-β finds the projection of v then BKZ-β finds v with "high probability" for large β.

The rest of this paper uses the inequality as a black box without regard to the rationale.

For the LPR PKE, the first $n+\kappa$ entries in the target vector $(a, \text{First}_\kappa(e), 1)$ are drawn independently and uniformly at random from χ. Each entry has square $\sum_i \chi_i i^2 = s^2$ on average (since χ has average 0 and standard deviation s), so the squared 2-norm of $(a, \text{First}_\kappa(e), 1)$ is $(n+\kappa)s^2+1$ on average. The standard heuristic analysis treats the squared 2-norm as being exactly its average, concluding for this PKE that BKZ-β works if and only if $((n+\kappa)s^2+1)^{1/2} < (d/\beta)^{1/2}\delta^{2\beta-d-1}q^{\kappa/d}$; i.e., if and only if $\text{StandardRatio}(n,q,s,\kappa,\beta) < 1$, with the notation of Definition 1.1.1.

Other PKEs do not necessarily choose (a,e) this way: consider, e.g., a PKE that chooses a as a fixed-weight ternary vector. To apply the standard analysis to such cases, the literature calculates s so that $(n+\kappa)s^2+1$ is a reasonable estimate of the squared 2-norm of $(a, \text{First}_\kappa(e), 1)$, and concludes heuristically that BKZ-β works if and only if $\text{StandardRatio}(n,q,s,\kappa,\beta) < 1$.

2.4 A known flaw in the standard analysis

Note that δ increases with β until β reaches 36, contrary to ample evidence that, e.g., BKZ-20 usually finds shorter vectors than BKZ-10 does. As an extreme case, if one takes $\beta = 2$ (or any $\beta < 13$), then $\delta < 1$, so the first heuristic says that BKZ-β finds a nonzero vector of length exponentially below $(\det L)^{1/d}$. In fact, for most lattices, such vectors do not even exist, so certainly BKZ-β will not find them.

The standard heuristic conclusion says that, for any particular (q,s), BKZ-2 breaks LPR for all n above an easily calculated bound. Choosing (q,s), calculating that bound, and simply trying BKZ-2 shows that, no, BKZ-2 does not in fact do this. The standard patch for this flaw is to simply disallow small values of β: for example, require $\beta \geq 60$.

For some of this paper's calculations, it suffices to assume $\beta \geq 2$, ensuring that the exponent $1/2(\beta-1)$ is defined. At many points in the logic, β/n is known to grow asymptotically as $Y_0 + o(1)$ for some positive real number Y_0, implying $\beta \geq 60$ for all sufficiently large n. However, Theorem 1.2.1(2) does not assume any particular asymptotic growth of β/n, and the conclusion of Theorem 1.2.1(2) would be incorrect if the hypothesis $\beta \geq 60$ were weakened to $\beta \geq 2$.

2.5 Asymptotics for the standard message-recovery attack

Consider now the problem of recovering the encryption secrets (b,d) from (G,B) where $B = Gb + d$. The standard analysis handles this exactly the same way as the key-recovery

attack, except for starting with the distribution of (b, d) rather than the distribution of (a, e).

In the case of the LPR PKE, the distributions are the same, so the conclusions are the same. The attacker will then prefer to carry out the key-recovery attack since it breaks many ciphertexts.

Presumably the use of key-recovery attacks as multi-ciphertext attacks is the motivation for, e.g., [**L6P11**, Section 6] saying "arguably, the secret key ought to be better-protected than any individual ciphertext". One can easily modify the LPR PKE to take a larger distribution for (a, e) than for (b, d), i.e., to use a larger error distribution for key generation than for encryption. This paper's second-order asymptotics make it easy to see the effect of, e.g., making errors 1 bit larger; see Section 1.3.

2.6 Further attacks

The standard analysis also considers the problem of recovering (b, c, d) given G, A, $B = Gb + d$, and $C = Ab + c$, i.e., from a public key and an encryption of 0. A successful recovery attack immediately gives an IND-CPA attack.

Structurally, this problem provides more "samples" to the attacker, allowing κ to be chosen as large as $2n$. This is covered by this paper's analysis: Theorem 1.2.1(2) ends up with $\kappa/n \in x_0 + o(1)$ with $x_0 \leq 1$, even if κ/n is initially allowed to be much larger than 1.

The situation would change if this paper allowed S_0 to be above $1/2$: the same optimizations would then produce $x_0 > 1$. For essentially the same reason, a close look at [**A3CDD+18**] finds, e.g., "Frodo-0640" listed as 2^{142} on [**A3CDD+18**, page 29] for "n LWE samples" but as only 2^{141} on [**A3CDD+18**, page 35] for "$2n$ LWE samples". This is a counterexample to, e.g., [**D4HKLS21**, page 3] saying "we believe that in practice the MLWE problem with k samples is *no easier* than with 1 sample" (emphasis added).

IND-CCA2 attacks against lattice KEMs can be much easier than the usual lattice attacks against the underlying PKEs. Examples include the Round2 break in [**B3DG20**] and a more recent attack exploiting derandomization in FrodoKEM. However, lattice attacks against PKEs seem to be the top threat for most lattice proposals.

3 Proof of the main theorem

As noted in Section 1, the paper includes two proofs of Theorem 1.2.1. For readers who simply want high assurance that the theorem is correct, the computer-verified proof (see Appendix A) is better. For readers who want to understand how to prove the theorem, the traditional-format proof in this section is better.

3.1 Supporting theorems

Some lemmas are factored out of the main theorem as the following separate theorems.

Theorem 3.1.1 ($o(1)$ lower bounds). *Let N be an infinite set of nonnegative integers. Let φ be a function from N to $\mathbb{R}$. Assume that $\{n \in N : \varphi(n) < -\epsilon\}$ is finite for each real number $\epsilon > 0$. Then there is a function ψ from N to $\mathbb{R}$ such that (1) $\varphi(n) \geq \psi(n)$ for all $n \in N$, (2) $\psi(n) \leq 0$ for all $n \in N$, and (3) $\psi(n) \to 0$ as $n \to \infty$.*

The proof technique is standard. Typically the conclusion would be written as "$\varphi(n) \geq o(1)$", meaning that $\varphi(n)$ is bounded below by *some* function converging to 0; this is not to be confused with the false statement that $\varphi(n)$ is bounded below by *every* function converging to 0. This paper avoids "$\geq o(1)$" notation.

Theorem 3.1.1 is used in the proof of Theorem 1.2.1(2).

Proof. The set $\{\varphi(n) : n \in N, \varphi(n) < -1\}$ is finite. Write ℓ for the minimum element of this set, or for -1 if the set is empty. Then $\varphi(n) \geq \ell$ for all $n \in N$.

(The first version of this paper had instead defined ℓ as the minimum element of $\{n \in N : \varphi(n) < -1\}$; thanks to "@ladygaladriel21" for pointing out the error.)

Define $S_n = \{0\} \cup \{\varphi(x) : x \in N, x \geq n\}$ for each $n \in N$. Then S_n is nonempty, and has a lower bound (namely ℓ), so it has a greatest lower bound in $\mathbb{R}$. Define $\psi(n)$ as this greatest lower bound. (This is the infimum of S_n.)

In particular, $\varphi(n) \in S_n$, so $\psi(n) \leq \varphi(n)$ as claimed. Also $0 \in S_n$, so $\psi(n) \leq 0$ as claimed.

If $m, n \in N$ have $n \geq m$ then $S_n \subseteq S_m$ so $\psi(n) \geq \psi(m)$. Hence ψ is a nondecreasing function. It has 0 as an upper bound, so it has a limit $L \leq 0$.

For each real number $\epsilon > 0$, the set $T = \{m \in N : \varphi(m) < -\epsilon\}$ is finite. Any $n \in N$ larger than all elements of T has $\varphi(x) \geq -\epsilon$ for all $x \geq n$, so S_n has $-\epsilon$ as a lower bound, so $\psi(n) \geq -\epsilon$, so $L \geq -\epsilon$. Hence $L = 0$ as claimed. □

Theorem 3.1.2 (first-order κ optimization). *Let Q_0, S_0 be real numbers with $0 \leq S_0 < 1/2 < Q_0 - S_0$. Then the quantity $(1+x)/(1-2S_0+2Q_0x/(1+x))$ for real numbers $x \geq 0$ has minimum value $2Q_0/(Q_0 - S_0 + 1/2)^2$, achieved uniquely for $x = (Q_0 + S_0 - 1/2)/(Q_0 - S_0 + 1/2)$.*

The proof is a calculus exercise. Theorem 3.1.2 is used in the proof of Theorem 1.2.1(2).

Proof. Note first that all denominators appearing here are positive: $1 + x \geq 1$ since $x \geq 0$; $Q_0 - S_0 + 1/2 > 1$ since $Q_0 - S_0 > 1/2$; $Q_0 > 1/2$ since $Q_0 - S_0 > 1/2$ and $S_0 \geq 0$; $1 - 2S_0 + 2Q_0x/(1+x) > 0$ since $1 - 2S_0 > 0$ and $2Q_0x \geq 0$.

Define $x_0 = (Q_0 + S_0 - 1/2)/(Q_0 - S_0 + 1/2)$. Then $1 + x_0 = 2Q_0/(Q_0 - S_0 + 1/2)$, so $2Q_0x_0/(1+x_0) = Q_0 + S_0 - 1/2$, so $1 - 2S_0 + 2Q_0x_0/(1+x_0) = Q_0 - S_0 + 1/2$, so $(1+x_0)/(1-2S_0+2Q_0x_0/(1+x_0)) = 2Q_0/(Q_0 - S_0 + 1/2)^2$.

Define $\varphi(x) = (1+x)/(1-2S_0+2Q_0x/(1+x))$. The derivative $\varphi'(x)$ is

$$\frac{(1+x)((Q_0 - S_0 + 1/2)x - (Q_0 + S_0 - 1/2))}{2((1/2 - S_0)(1+x) + Q_0x)^2},$$

which is 0 for $x = x_0$, negative for $x < x_0$, and positive for $x > x_0$. Hence $\varphi(x)$ achieves its minimum value uniquely at $x = x_0$. This minimum is $\varphi(x_0) = 2Q_0/(Q_0 - S_0 + 1/2)^2$. □

Theorem 3.1.3 (main-term κ optimization). *Let $n, q, \kappa, \beta, \delta$ be real numbers with $2 \leq n$, $2 \leq q$, $1 \leq \kappa$, $2 \leq \beta$, and $1 < \delta$. Then*

$$(2\beta - n - \kappa - 2)\lg\delta + \frac{\kappa}{n+\kappa+1}\lg q \leq (2\beta - 1)\lg\delta + \lg q - 2\sqrt{(n+1)(\lg\delta)\lg q}.$$

The proof is another calculus exercise. Theorem 3.1.3 is used in the proof of Theorem 1.2.1(2).

Proof. Define $x_0 = \sqrt{(n+1)(\lg q)/\lg\delta} - n - 1$. Note that $(n + x_0 + 1)^2 \lg\delta = (n+1)\lg q$, and $(n + x_0 + 1)\lg\delta = \sqrt{(n+1)(\lg\delta)\lg q} = ((n+1)/(n+x_0+1))\lg q$.

Define $\varphi(x) = (n+x+1)\lg\delta + ((n+1)/(n+x+1))\lg q$ for each real number $x \geq 0$. The derivative $\varphi'(x)$ is $\lg\delta - (n+1)(n+x+1)^{-2}\lg q$, which is 0 for $x = x_0$, negative for $x < x_0$, and positive for $x > x_0$. Hence $\varphi(x) \geq \varphi(x_0) = (n+x_0+1)\lg\delta + ((n+1)/(n+x_0+1))\lg q = 2\sqrt{(n+1)(\lg\delta)\lg q}$.

In particular,

$$\begin{aligned}&(2\beta-n-\kappa-2)\lg\delta-\frac{\kappa}{n+\kappa+1}\lg q\\&\quad=(2\beta-1)\lg\delta+\lg q-(n+\kappa+1)\lg\delta-\frac{n+1}{n+\kappa+1}\lg q\\&\quad=(2\beta-1)\lg\delta+\lg q-\varphi(\kappa)\\&\quad\le(2\beta-1)\lg\delta+\lg q-2\sqrt{(n+1)(\lg\delta)\lg q}\end{aligned}$$

as claimed. □

Theorem 3.1.4 (monotonicity in the block size). *Let n,q,s,κ,x be real numbers such that $2\le n$; $2\le q$; $0<s$; $1\le\kappa$; and $60\le x$. Define $\delta=(x(\pi x)^{1/x}/(2\pi\exp 1))^{1/2(x-1)}$. Then $\delta>1$, and $\mathrm{StandardRatio}(n,q,s,\kappa,x)\ge\mathrm{StandardRatio}(n,q,s,\kappa,y)$ for all real numbers $y\ge x$.*

The proof is a series of calculus exercises. Theorem 3.1.4 is used in the proof of Theorem 1.2.1(2).

Proof. First note that $2(x-1)\log\delta=\log x+(1/x)\log\pi x-\log(2\pi\exp 1)$. Now $\log x\ge\log 60>4>1+\log(2\pi\exp 1)$ so $2(x-1)\log\delta>1+(1/x)\log\pi x>1$ so $2\log\delta>1/(x-1)$. In particular, $\log\delta>0$, so $\delta>1$.

Write $\rho=\mathrm{StandardRatio}(n,q,s,\kappa,x)$. The main point of the proof is that the partial derivative ρ' of ρ with respect to x (i.e., the derivative of $x\mapsto\rho$ for fixed n,q,s,κ) is negative for all $x\ge 60$.

First $2(x-1)\delta'/\delta+2\log\delta=1/x+1/x^2-(1/x^2)\log\pi x$, where δ' means the derivative of δ with respect to x. Also $x^2-1<x^2$ so $1/x+1/x^2<1/(x-1)$ so $2(x-1)\delta'/\delta+2\log\delta<1/(x-1)-(1/x^2)\log\pi x<1/(x-1)$. Hence $2(x-1)\delta'/\delta<0$, implying $\delta'<0$.

By definition $\rho=((n+\kappa)s^2+1)^{1/2}/(d/x)^{1/2}\delta^{2x-d-1}q^{\kappa/d}$ where $d=n+\kappa+1$. Hence $\log\rho=\cdots+(1/2)\log x-(2x-d-1)\log\delta$ where $\cdots$ is independent of x, so $\rho'/\rho=1/2x-(2x-d-1)\delta'/\delta-2\log\delta$. Substitute $2\log\delta=1/x+1/x^2-(1/x^2)\log\pi x-(2x-2)\delta'/\delta$ to see that $\rho'/\rho=(d-1)\delta'/\delta-1/2x-1/x^2+(1/x^2)\log\pi x$.

Note that $\log\pi x-x/2<0$ (since $\log(60\pi)<30$ and the derivative $1/x-1/2$ is negative), so $-1/2x+(1/x^2)\log\pi x<0$, so $\rho'/\rho<(d-1)\delta'/\delta-1/x^2$. Also, $d-1>0$ since $d=n+\kappa+1\ge 4$, so $(d-1)\delta'/\delta<0$. Hence $\rho'<0$ as claimed. Consequently replacing x with a larger value decreases ρ as claimed. □

Theorem 3.1.5 (first-order asymptotics of the standard ratio). *Let Q_0,S_0,X_0,Y_0 be real numbers such that $0<Q_0$; $-1/2<S_0$; $0\le X_0$; $0<Y_0$; and $0<1-2S_0+2Q_0X_0/(1+X_0)$. Define*

$$Z_0=\frac{1+X_0}{1-2S_0+2Q_0X_0/(1+X_0)}.$$

Let N be an infinite subset of $\{2,3,4,5,\ldots\}$. Let $n\mapsto q$, $n\mapsto s$, $n\mapsto\kappa$, $n\mapsto\beta$ be functions from N to $\mathbb{R}$ such that

$$\begin{aligned}2\le q,&\qquad\lg q\in(Q_0+o(1))\lg n,\\0<s,&\qquad\lg s\in(S_0+o(1))\lg n,\\1\le\kappa,&\qquad\kappa/n\in X_0+o(1),\\2\le\beta,&\qquad\beta/n\in Y_0+o(1),\end{aligned}$$

(0) One has

$$\frac{2\lg\mathrm{StandardRatio}(n,q,s,\kappa,\beta)}{\lg n}\in 2S_0-1+\frac{1+X_0}{Y_0}-\frac{2Q_0X_0}{1+X_0}+o(1).$$

(1) If $Y_0>Z_0$ then $\mathrm{StandardRatio}(n,q,s,\kappa,\beta)<1$ for all sufficiently large n. (2) If $Y_0<Z_0$ then $\mathrm{StandardRatio}(n,q,s,\kappa,\beta)>1$ for all sufficiently large n.

Parts (1) and (2) follow easily from part (0). The proof of part (0) starts with the hypothesized asymptotics for q, s, κ, β and computes asymptotics for $q^{\kappa/d}$, d/β, etc., culminating in $\operatorname{StandardRatio}(n, q, s, \kappa, \beta)$.

Part (1) is used in the proof of Theorem 1.2.1(2). Part (2) is not used directly but shows that the cutoff in part (1) is tight. This theorem is also a warmup for Theorem 3.1.6, which draws more precise conclusions from more precise hypotheses.

Proof. Write $\rho = \operatorname{StandardRatio}(n, q, s, \kappa, \beta)$. By definition

$$\rho = \frac{((n+\kappa)s^2+1)^{1/2}}{(d/\beta)^{1/2}\delta^{2\beta-d-1}q^{\kappa/d}}$$

where $d = n+\kappa+1$ and $\delta = (\beta(\pi\beta)^{1/\beta}/(2\pi\exp 1))^{1/2(\beta-1)}$. Now $d/n = 1+\kappa/n+1/n \in 1+X_0+o(1)$, so $n/d \in 1/(1+X_0)+o(1)$, so $\kappa/d \in X_0/(1+X_0)+o(1)$. Meanwhile $(\lg q)/\lg n \in Q_0+o(1)$, so $(\kappa \lg q)/(d \lg n) \in Q_0X_0/(1+X_0)+o(1)$.

Also $d/\beta \in (1+X_0)/Y_0+o(1)$ since $\beta/n \in Y_0+o(1)$. What matters for this proof is the weaker statement $d/\beta \in n^{o(1)}$, i.e., $(\lg(d/\beta))/\lg n \in o(1)$.

Next $s^2 \in n^{2S_0+o(1)}$ and $n+\kappa \in n^{1+o(1)}$ so $(n+\kappa)s^2 \in n^{1+2S_0+o(1)}$. The exponent $1+2S_0$ is positive, so also $(n+\kappa)s^2+1 \in n^{1+2S_0+o(1)}$, so

$$\frac{\lg((n+\kappa)s^2+1)}{\lg n} \in 1+2S_0+o(1).$$

By hypothesis $\beta \in (Y_0+o(1))n$ with $Y_0 > 0$, so $\beta \in \Theta(n)$, so $\lg \pi\beta \in O(\lg n)$, so $(\lg \pi\beta)/\beta \in O(\lg n)/n$. What matters here is merely that $(\lg \pi\beta)/\beta \in o(1)$, so

$$2(\beta-1)\lg\delta = \lg\beta + \frac{\lg \pi\beta}{\beta} - \lg(2\pi\exp 1) \in \lg n + \lg Y_0 - \lg(2\pi\exp 1) + o(1).$$

In particular, $2(\beta-1)\lg\delta \in (1+o(1))\lg n$.

Divide $(d-1)/n \in 1+X_0+o(1)$ by $(\beta-1)/n \in Y_0+o(1)$ to see that $(d-1)/(\beta-1) \in (1+X_0)/Y_0+o(1)$. Consequently $(2\beta-d-1)/2(\beta-1) \in 1-(1+X_0)/2Y_0+o(1)$. Multiply by $2(\beta-1)\lg\delta \in (1+o(1))\lg n$ to see that $((2\beta-d-1)\lg\delta)/\lg n \in 1-(1+X_0)/2Y_0+o(1)$.

Now add to see that

$$\begin{aligned}\frac{2\lg\rho}{\lg n} &= \frac{\lg((n+\kappa)s^2+1)}{\lg n} - \frac{\lg(d/\beta)}{\lg n} - \frac{2\lg\delta^{2\beta-d-1}}{\lg n} - \frac{2(\kappa/d)\lg q}{\lg n} \\ &\in 2S_0 - 1 + \frac{1+X_0}{Y_0} - \frac{2Q_0X_0}{1+X_0} + o(1)\end{aligned}$$

as claimed.

If $Y_0 > Z_0$, i.e., $Y_0 > (1+X_0)/(1-2S_0+2Q_0X_0/(1+X_0))$, then the limit $1+2S_0-2+(1+X_0)/Y_0-2Q_0X_0/(1+X_0)$ of $(2\lg\rho)/\lg n$ is negative, so $(2\lg\rho)/\lg n < 0$ for all sufficiently large n, so $\rho < 1$ for all sufficiently large n as claimed.

If $Y_0 < Z_0$ then the limit is positive, so $\rho > 1$ for all sufficiently large n as claimed. □

Theorem 3.1.6 (second-order asymptotics of the standard ratio). *Let $Q_0, Q_1, S_0, S_1, X_0, X_1, Y_0, Y_1$ be real numbers such that $0 < Q_0$; $-1/2 < S_0$; $(0,0) \le (X_0, X_1) \le (1,0)$; $0 < Y_0$; $(Y_0, Y_1) \le (1+X_0, X_1)$; and $0 < 1-2S_0+2Q_0X_0/(1+X_0)$. Define*

$$Z_0 = \frac{1+X_0}{1-2S_0+2Q_0X_0/(1+X_0)},$$

$$Z_1 = \left(2S_1+\lg Z_0+\frac{X_1}{Z_0}-\left(2-\frac{1+X_0}{Z_0}\right)\lg\frac{Z_0}{2\pi\exp 1}-\frac{2Q_1X_0}{1+X_0}-\frac{2Q_0X_1}{(1+X_0)^2}\right)\frac{Z_0^2}{1+X_0}.$$

Let N be an infinite subset of $\{2,3,4,5,\ldots\}$. Let $n \mapsto q$, $n \mapsto s$, $n \mapsto \kappa$, $n \mapsto \beta$ be functions from N to $\mathbb{R}$ such that

$$\begin{aligned} &2 \le q, && \lg q \in Q_0 \lg n + Q_1 + o(1), \\ &0 < s, && \lg s \in S_0 \lg n + S_1 + o(1), \\ &1 \le \kappa \le n, && \kappa/n \in X_0 + (X_1 + o(1))/\lg n, \\ &2 \le \beta \le n + \kappa + 1, && \beta/n \in Y_0 + (Y_1 + o(1))/\lg n. \end{aligned}$$

(1) If $(Y_0, Y_1) > (Z_0, Z_1)$ then StandardRatio$(n, q, s, \kappa, \beta) < 1$ *for all sufficiently large n.*
(2) If $(Y_0, Y_1) < (Z_0, Z_1)$ then StandardRatio$(n, q, s, \kappa, \beta) > 1$ *for all sufficiently large n.*

The proof follows the same lines as Theorem 3.1.5 but involves more calculations to track everything to second order rather than just first order.

Part (1) is used in the proof of Theorem 1.2.1(1). Part (2) is not used directly but shows that the cutoff in part (1) is tight.

Proof. Write $\rho =$ StandardRatio(n, q, s, κ, β). By definition

$$\rho = \frac{((n+\kappa)s^2 + 1)^{1/2}}{(d/\beta)^{1/2} \delta^{2\beta - d - 1} q^{\kappa/d}}$$

where $d = n + \kappa + 1$ and $\delta = (\beta(\pi\beta)^{1/\beta}/(2\pi \exp 1))^{1/2(\beta-1)}$.

Note that arithmetic on quantities of the form $A_0 + (A_1 + o(1))/\lg n$, where $A_0, A_1 \in \mathbb{R}$, matches arithmetic on elements $A_0 + A_1\epsilon$ of the ring $\mathbb{R}[\epsilon]/\epsilon^2$. For example,

$$\left(A_0 + \frac{A_1 + o(1)}{\lg n}\right)\left(B_0 + \frac{B_1 + o(1)}{\lg n}\right) \subseteq A_0 B_0 + \frac{A_0 B_1 + A_1 B_0 + o(1)}{\lg n}$$

since $A_0 o(1)/\lg n$, $A_1 o(1)/\lg n$, and $(A_1 + o(1))(B_1 + o(1))/(\lg n)^2$ are all contained in $o(1)/\lg n$. As another example, $1/(A_0 + (A_1 + o(1))/\lg n)$ is $1/A_0 + (-A_1/A_0^2 + o(1))/\lg n$ if $A_0 \neq 0$.

In particular, $d/n = 1 + \kappa/n + 1/n \in 1 + X_0 + (X_1 + o(1))/\lg n$, so

$$\frac{n}{d} \in \frac{1}{1+X_0} + \frac{-X_1/(1+X_0)^2 + o(1)}{\lg n},$$

so

$$\begin{aligned} \frac{\kappa}{d} &\in \frac{X_0}{1+X_0} + \frac{X_1/(1+X_0) - X_0 X_1/(1+X_0)^2 + o(1)}{\lg n} \\ &= \frac{X_0}{1+X_0} + \frac{X_1/(1+X_0)^2 + o(1)}{\lg n}. \end{aligned}$$

Meanwhile $(\lg q)/\lg n \in Q_0 + (Q_1 + o(1))/\lg n$. Multiply to see that

$$\frac{\kappa \lg q}{d \lg n} \in \frac{Q_0 X_0}{1+X_0} + \frac{Q_1 X_0/(1+X_0) + Q_0 X_1/(1+X_0)^2 + o(1)}{\lg n}.$$

The proof does not need as much precision for d/β: one has $n/\beta \in 1/Y_0 + o(1)$ and $d/n \in 1 + X_0 + o(1)$, so $d/\beta \in (1+X_0)/Y_0 + o(1)$, so

$$\frac{\lg(d/\beta)}{\lg n} \in \frac{\lg(1+X_0) - \lg Y_0 + o(1)}{\lg n}.$$

Next $s^2 \in n^{2S_0} 2^{2S_1 + o(1)} = n^{2S_0}(2^{2S_1} + o(1))$ and $n + \kappa \in n(1 + X_0 + o(1))$ so $(n+\kappa)s^2 \in n^{1+2S_0}((1+X_0)2^{2S_1} + o(1))$. The exponent $1 + 2S_0$ is positive, so also $(n+\kappa)s^2 + 1 \in n^{1+2S_0}((1+X_0)2^{2S_1} + o(1))$. Hence

$$\frac{\lg((n+\kappa)s^2 + 1)}{\lg n} \in 1 + 2S_0 + \frac{\lg(1+X_0) + 2S_1 + o(1)}{\lg n}.$$

By hypothesis $\beta \in (Y_0 + o(1))n$ with $Y_0 > 0$, so $\beta \in \Theta(n)$, so $\lg \pi\beta \in O(\lg n)$, so $(\lg \pi\beta)/\beta \in O(\lg n)/n$. What matters here is merely that $(\lg \pi\beta)/\beta \in o(1)$, so

$$2(\beta-1)\lg\delta = \lg\beta + \frac{\lg\pi\beta}{\beta} - \lg(2\pi\exp 1) \in \lg n + \lg Y_0 - \lg(2\pi\exp 1) + o(1).$$

Divide $(d-1)/n \in 1 + X_0 + (X_1 + o(1))/\lg n$ by $(\beta-1)/n \in Y_0 + (Y_1 + o(1))/\lg n$ to see that

$$\frac{d-1}{\beta-1} \in \frac{1+X_0}{Y_0} + \frac{X_1/Y_0 - (1+X_0)Y_1/Y_0^2 + o(1)}{\lg n}.$$

Consequently

$$\frac{2\beta-d-1}{2(\beta-1)} = 1 - \frac{d-1}{2(\beta-1)} \in 1 - \frac{1+X_0}{2Y_0} - \frac{X_1/2Y_0 - (1+X_0)Y_1/2Y_0^2 + o(1)}{\lg n}.$$

Multiply by $2(\beta-1)(\lg\delta)/\lg n \in 1 + (\lg Y_0 - \lg(2\pi\exp 1) + o(1))/\lg n$ to see that

$$\begin{aligned}\frac{(2\beta-d-1)\lg\delta}{\lg n} &\in 1 - \frac{1+X_0}{2Y_0} \\ &+ \frac{-X_1/2Y_0 + (1+X_0)Y_1/2Y_0^2 + (1-(1+X_0)/2Y_0)(\lg Y_0 - \lg(2\pi\exp 1)) + o(1)}{\lg n}.\end{aligned}$$

Now add to see that

$$\begin{aligned}\frac{2\lg\rho}{\lg n} &= \frac{\lg((n+\kappa)s^2+1)}{\lg n} - \frac{\lg(d/\beta)}{\lg n} - \frac{2\lg\delta^{2\beta-d-1}}{\lg n} - \frac{2(\kappa/d)\lg q}{\lg n} \\ &\in 1 + 2S_0 + \frac{\lg(1+X_0)+2S_1}{\lg n} - \frac{\lg(1+X_0)-\lg Y_0}{\lg n} - 2 + \frac{1+X_0}{Y_0} \\ &\quad + \frac{X_1/Y_0 - (1+X_0)Y_1/Y_0^2 - (2-(1+X_0)/Y_0)(\lg Y_0 - \lg(2\pi\exp 1))}{\lg n} \\ &\quad - \frac{2Q_0X_0}{1+X_0} - \frac{2Q_1X_0/(1+X_0) + 2Q_0X_1/(1+X_0)^2}{\lg n} + \frac{o(1)}{\lg n}.\end{aligned}$$

If $Y_0 > Z_0$, i.e., $Y_0 > (1+X_0)/(1-2S_0+2Q_0X_0/(1+X_0))$, then the limit $1 + 2S_0 - 2 + (1+X_0)/Y_0 - 2Q_0X_0/(1+X_0)$ of $(2\lg\rho)/\lg n$ is negative, so $(2\lg\rho)/\lg n < 0$ for all sufficiently large n, so $\rho < 1$ for all sufficiently large n as claimed.

If $Y_0 < Z_0$ then the limit is positive, so $\rho > 1$ for all sufficiently large n as claimed.

(The previous two paragraphs are identical to paragraphs in the proof of Theorem 3.1.5, and can be replaced by an application of Theorem 3.1.5, but the real work in this proof comes from handling the second-order asymptotics needed for the case $Y_0 = Z_0$.)

Assume from now on that $Y_0 = Z_0$. Then the limit is 0, and one has $(2\lg\rho)/\lg n \in (\Delta + o(1))/\lg n$ where

$$\begin{aligned}\Delta &= 2S_1 + \lg Z_0 + \frac{X_1}{Z_0} - \frac{(1+X_0)Y_1}{Z_0^2} \\ &\quad - \left(2 - \frac{1+X_0}{Z_0}\right)\lg\frac{Z_0}{2\pi\exp 1} - \frac{2Q_1X_0}{1+X_0} - \frac{2Q_0X_1}{(1+X_0)^2} \\ &= \frac{(1+X_0)(Z_1-Y_1)}{Z_0^2};\end{aligned}$$

note that the $\lg(1+X_0)$ terms cancel.

If $Y_1 > Z_1$ then $\Delta < 0$ so $\rho < 1$ for all sufficiently large n as claimed. If $Y_1 < Z_1$ then $\Delta > 0$ so $\rho > 1$ for all sufficiently large n as claimed. The only remaining possibility is $(Y_0, Y_1) = (Z_0, Z_1)$, and the theorem statement makes no claims regarding this case. □

3.2 The main theorem

The main proof relies on the theorems from Section 3.1.

Theorem 1.2.1 (asymptotic growth of the standard block size). *Let Q_0, Q_1, S_0, S_1 be real numbers such that $0 \le S_0 \le 1/2 < Q_0 - S_0$. Let N be an infinite subset of $\{2, 3, 4, 5, \ldots\}$. Let $n \mapsto q$ and $n \mapsto s$ be functions from N to $\mathbb{R}$ such that*

$$\begin{aligned} 2 \le q, &\qquad \lg q \in Q_0 \lg n + Q_1 + o(1), \\ 0 < s, &\qquad \lg s \in S_0 \lg n + S_1 + o(1). \end{aligned}$$

Define $x_0 = (Q_0 + S_0 - 1/2)/(Q_0 - S_0 + 1/2)$; $z_0 = 2Q_0/(Q_0 - S_0 + 1/2)^2$; and

$$z_1 = \left(2S_1 + \lg z_0 - \left(S_0 - Q_0 + \frac{3}{2}\right) \lg \frac{z_0}{2\pi \exp 1} - \frac{Q_1(Q_0 + S_0 - \frac{1}{2})}{Q_0}\right) \frac{2Q_0}{(Q_0 - S_0 + \frac{1}{2})^3}.$$

(1) There are functions $n \mapsto \kappa$ and $n \mapsto \beta$ from N to $\mathbb{Z}$ such that

$$\begin{aligned} &1 \le \kappa \le n \text{ for all } n, && \kappa/n \in x_0 + o(1)/\lg n, \\ &2 \le \beta \le n + \kappa + 1 \text{ for all } n, && \beta/n \in z_0 + (z_1 + o(1))/\lg n, \qquad \text{and} \\ &\text{StandardRatio}(n, q, s, \kappa, \beta) < 1 \text{ for all sufficiently large } n. \end{aligned}$$

(2) Let $n \mapsto \kappa$ and $n \mapsto \beta$ be functions from N to $\mathbb{R}$ such that

$$\begin{aligned} &1 \le \kappa \le 100n \text{ for all } n, \\ &60 \le \beta \le n + \kappa + 1 \text{ for all } n, \qquad \text{and} \\ &\text{StandardRatio}(n, q, s, \kappa, \beta) \le 1 \text{ for all sufficiently large } n. \end{aligned}$$

Then $\beta \ge \ell$ for some function $n \mapsto \ell$ with $\ell/n \in z_0 + (z_1 + o(1))/\lg n$.

Conceptually, seeing the formulas for x_0, z_0, z_1 is a calculus exercise starting from Theorem 3.1.6. Choosing X_0 to minimize Z_0 produces $X_0 = x_0$ and $Z_0 = z_0$; it also produces $Z_1 = z_1$, independently of X_1, since the X_1 terms in Z_1 cancel when $X_0 = x_0$.

Proving Theorem 1.2.1(1) takes more work to glue together choices of β where (Y_0, Y_1) is *larger* than (z_0, z_1) into a choice of β where (Y_0, Y_1) is *equal* to (z_0, z_1).

Proving Theorem 1.2.1(2) takes more work to deal with the fact that κ and β are not assumed to have any particular asymptotic growth rate. Hypothesizing growth rates as in Theorem 3.1.6 would allow a simpler proof using Theorem 3.1.6(2) in place of Theorems 3.1.1, 3.1.2, 3.1.3, and 3.1.4; but this would not logically rule out the possibility that a different growth rate allows smaller β.

Proof. Note first the following equations and inequalities:

- $Q_0 - S_0 + 1/2 > 1$;
- $0 < x_0$ since $1 < 2Q_0 + 2S_0$;
- $x_0 \le 1$ since $1 - x_0 = 2(1/2 - S_0)/(Q_0 - S_0 + 1/2) \ge 0$;
- $1 + x_0 = 2Q_0/(Q_0 - S_0 + 1/2)$ so $2Q_0x_0/(1 + x_0) = Q_0 + S_0 - 1/2$ so $1 - 2S_0 + 2Q_0x_0/(1 + x_0) = Q_0 - S_0 + 1/2 > 1$;
- $0 < z_0$ since $0 < Q_0$; and
- $z_0 = (1 + x_0)/(Q_0 - S_0 + 1/2)$ so $z_0 < 1 + x_0$.

Proof of part (1). Define a function $n \mapsto \kappa$ from N to $\mathbb{Z}$ by the formula $\kappa = \lceil x_0 n \rceil$.

Note that $0 < x_0 n \le n$ since $0 < x_0 \le 1$, so $1 \le \kappa \le n$. Also $x_0 n \le \kappa < x_0 n + 1$ so $x_0 \le \kappa/n < x_0 + 1/n$ so $\kappa/n \in x_0 + o(1)/\lg n$.

Next, for each integer $j \ge 1$, define a function $n \mapsto \beta_j$ from N to $\mathbb{Z}$ by the formula $\beta_j = \min\left\{n + \kappa + 1, \max\left\{2, \left\lceil (z_0 + (z_1 + 2^{-j})/\lg n)n \right\rceil\right\}\right\}$. By construction $2 \le \beta_j \le n + \kappa + 1$ since $n + \kappa + 1 > 2$. Also note for future reference that $\beta_1 \ge \beta_2 \ge \beta_3 \ge \cdots$.

Abbreviate $(z_1 + 2^{-j})/\lg n$ as ϵ. One has $\epsilon \in o(1)$. In particular, $-z_0/2 < \epsilon$ for all sufficiently large n, since $-z_0/2 < 0$. Hence $z_0/2 < z_0 + \epsilon$ for all sufficiently large n, so $2 \le (z_0/2)n < (z_0 + \epsilon)n \le \lceil (z_0 + \epsilon)n \rceil$ for all sufficiently large n, so $\max\{2, \lceil (z_0 + \epsilon)n \rceil\} = \lceil (z_0 + \epsilon)n \rceil$ for all sufficiently large n.

Similarly, $\epsilon \le 1 + x_0 - z_0$ for all sufficiently large n, since $1 + x_0 - z_0 > 0$. Hence $z_0 + \epsilon \le 1 + x_0$ for all sufficiently large n, so $(z_0 + \epsilon)n \le (1 + x_0)n = n + x_0 n \le n + \kappa$ for all sufficiently large n, so $\lceil (z_0 + \epsilon)n \rceil \le n + \kappa + 1$ for all sufficiently large n, so $\min\{n + \kappa + 1, \lceil (z_0 + \epsilon)n \rceil\} = \lceil (z_0 + \epsilon)n \rceil$ for all sufficiently large n.

Consequently $\beta_j = \left\lceil (z_0 + (z_1 + 2^{-j})/\lg n)n \right\rceil$ for all sufficiently large n; so $\beta_j/n \in z_0 + (z_1 + 2^{-j} + o(1))/\lg n$.

Check the hypotheses of Theorem 3.1.6(1), with $X_0 = x_0$, $X_1 = 0$, $Y_0 = z_0$, $Y_1 = z_1 + 2^{-j}$, $Z_0 = z_0$, and $Z_1 = z_1$, and β_j in place of β:

- $0 < Q_0$: indeed, $Q_0 = Q_0 - S_0 + S_0 > 1/2$.
- $-1/2 < S_0$: indeed, $0 \le S_0$.
- $(0, 0) \le (X_0, X_1)$: indeed, $0 < x_0$.
- $0 < Y_0$: indeed, $0 < z_0$.
- $(Y_0, Y_1) \le (1 + X_0, X_1)$: indeed, $z_0 < 1 + x_0$.
- $0 < 1 - 2S_0 + 2Q_0 X_0/(1 + X_0)$: indeed, $1 < 1 - 2S_0 + 2Q_0 x_0/(1 + x_0)$.
- $Z_0 = (1 + X_0)/(1 - 2S_0 + 2Q_0 X_0/(1 + X_0))$: indeed, $1 - 2S_0 + 2Q_0 x_0/(1 + x_0) = Q_0 - S_0 + 1/2$, so $(1 + x_0)/(1 - 2S_0 + 2Q_0 x_0/(1 + x_0)) = (1 + x_0)/(Q_0 - S_0 + 1/2) = 2Q_0/(Q_0 - S_0 + 1/2)^2 = z_0 = Z_0$.
- Z_1 formula: see below.
- N is an infinite subset of $\{2, 3, 4, 5 \ldots\}$: by hypothesis.
- $n \mapsto q$, $n \mapsto s$, $n \mapsto \kappa$, $n \mapsto \beta_j$ are functions from N to $\mathbb{R}$: indeed, they are functions from N to $\mathbb{Z}$.
- $2 \le q$: by hypothesis.
- $\lg q \in Q_0 \lg n + Q_1 + o(1)$: by hypothesis.
- $0 < s$: by hypothesis.
- $\lg s \in S_0 \lg n + S_1 + o(1)$: by hypothesis.
- $1 \le \kappa \le n$: proven above.
- $\kappa/n \in X_0 + (X_1 + o(1))/\lg n$: indeed, $\kappa/n \in x_0 + o(1)/\lg n$.
- $2 \le \beta_j \le n + \kappa + 1$: by construction.
- $\beta/n \in Y_0 + (Y_1 + o(1))/\lg n$: indeed, $\beta/n \in z_0 + (z_1 + 2^{-j})/\lg n$.
- $(Y_0, Y_1) > (Z_0, Z_1)$: indeed, $Y_0 = z_0 = Z_0$ and $Y_1 = z_1 + 2^{-j} > z_1 = Z_1$.

To check the formula

$$Z_1 = \left(2S_1 + \lg Z_0 + \frac{X_1}{Z_0} - \left(2 - \frac{1+X_0}{Z_0}\right) \lg \frac{Z_0}{2\pi \exp 1} - \frac{2Q_1X_0}{1+X_0} - \frac{2Q_0X_1}{(1+X_0)^2}\right) \frac{Z_0^2}{1+X_0},$$

observe first that the terms X_1/Z_0 and $2Q_0X_1/(1+X_0)^2$ disappear since $X_1 = 0$. The coefficient $2 - (1+X_0)/Z_0$ is $2 - (1+x_0)/z_0 = 2 - (Q_0 - S_0 + 1/2) = S_0 - Q_0 + 3/2$; the term $2Q_1X_0/(1+X_0)$ is $2Q_1x_0/(1+x_0) = (Q_1/Q_0)(Q_0 + S_0 - 1/2)$; and the factor $Z_0^2/(1+X_0)$ is $z_0^2/(1+x_0) = (1+x_0)/(Q_0 - S_0 + 1/2)^2 = 2Q_0/(Q_0 - S_0 + 1/2)^3$.

Consequently StandardRatio$(n, q, s, \kappa, \beta_j) < 1$ for all sufficiently large n by Theorem 3.1.6.

Define elements $m_1, m_2, \ldots$ of N as follows: m_j is the minimum element of N such that

- $m_j > m_i$ for all $i < j$ and
- StandardRatio$(n, q, s, \kappa, \beta_j) < 1$ for all $n \geq m_j$.

Then $m_1 < m_2 < \cdots$.

Define a function $n \mapsto \beta$ from N to $\mathbb{Z}$ as follows: $\beta = \beta_1$ if $n < m_2$; $\beta = \beta_2$ if $m_2 \leq n < m_3$; $\beta = \beta_3$ if $m_3 \leq n < m_4$; etc.

The point of this construction is that StandardRatio$(n, q, s, \kappa, \beta) < 1$ for all $n \geq m_1$. Indeed, there is some j for which $m_j \leq n < m_{j+1}$, so $\beta = \beta_j$, but StandardRatio$(n, q, s, \kappa, \beta_j) < 1$ for $n \geq m_j$ by definition of m_j.

All that remains is to see that β has the desired sizes: first, $2 \leq \beta \leq n + \kappa + 1$; second, $\beta/n \in z_0 + (z_1 + o(1))/\lg n$.

The first part is easy: one has $2 \leq \beta_j \leq n + \kappa + 1$ for each j, so $2 \leq \beta \leq n + \kappa + 1$.

For the second part, start with $\lceil (z_0 + (z_1 + 2^{-j})/\lg n)n \rceil \geq (z_0 + z_1/\lg n)n$, implying $\beta_j \geq \min\{n + \kappa + 1, \max\{2, (z_0 + z_1/\lg n)n\}\}$ for all j, implying $\beta \geq \min\{n + \kappa + 1, \max\{2, (z_0 + z_1/\lg n)n\}\}$, implying $\beta \geq (z_0 + z_1/\lg n)n$ for all sufficiently large n.

For an upper bound, note that $\beta \in \{\beta_j, \beta_{j+1}, \ldots\}$ for all $n \geq m_j$. One has $\beta_1 \geq \beta_2 \geq \beta_3 \geq \cdots$, so $\beta \leq \beta_j$ for all $n \geq m_j$. One has $\beta_j \leq 1 + (z_0 + (z_1 + 2^{-j})/\lg n)n$ for all sufficiently large n, so $\beta \leq 1 + (z_0 + (z_1 + 2^{-j})/\lg n)n$ for all sufficiently large n.

In other words, $(\beta/n - z_0)\lg n$ is at least z_1 for all sufficiently large n, and is at most $z_1 + 2^{-j} + (\lg n)/n$ for all sufficiently large n. Hence it converges to 0; i.e., $\beta/n \in z_0 + (z_1 + o(1))/\lg n$.

Proof of part (2). This part shows that $z_0 + (z_1 + o(1))/\lg n$ is minimal: at this level of detail of the asymptotics, no choice of functions $n \mapsto \kappa$, $n \mapsto \beta$ can do better than the functions constructed in part (1).

Suppose that the set $N' = \{n \in N : \beta/n \leq z_0 + (z_1 - \epsilon)/\lg n\}$ is infinite for some real number $\epsilon > 0$.

Write $B = (z_0 + (z_1 - \epsilon)/\lg n)n$. Then $\beta \leq B$ for all $n \in N'$, so StandardRatio$(n, q, s, \kappa, B) \leq$ StandardRatio(n, q, s, κ, β) for all $n \in N'$ by Theorem 3.1.4.

Define $N'' = \{n \in N' :$ StandardRatio$(n, q, s, \kappa, \beta) \leq 1\}$. Then N'' is infinite, since by assumption StandardRatio$(n, q, s, \kappa, \beta) \leq 1$ for all sufficiently large $n \in N$.

Now StandardRatio$(n, q, s, \kappa, B) \leq$ StandardRatio$(n, q, s, \kappa, \beta) \leq 1$ for all $n \in N''$. In other words, $((n+\kappa)s^2 + 1)^{1/2} \leq (d/B)^{1/2} D^{2B-d-1} q^{\kappa/d}$ for all $n \in N''$, where $d = n + \kappa + 1$ and $D = (B(\pi B)^{1/B}/(2\pi \exp 1))^{1/2(B-1)}$. Note that $D > 1$ by Theorem 3.1.4.

All values κ/n are in the compact interval $[0, 100]$. By Bolzano's theorem, there is some $X_0 \in [0, 100]$ and some infinite subset $N''' \subseteq N''$ such that κ/n for $n \in N'''$ converges to X_0, i.e., $\kappa/n \in X_0 + o(1)$.

Now apply Theorem 3.1.5(0), with (β, Y_0, δ, N) replaced by (B, z_0, D, N'''), to see that

$$\frac{2 \lg \text{StandardRatio}(n, q, s, \kappa, B)}{\lg n} = 2S_0 - 1 + \frac{1 + X_0}{z_0} - \frac{2Q_0X_0}{1 + X_0} + o(1).$$

The left side is ≤ 0, so $2S_0 - 1 + (1 + X_0)/z_0 - 2Q_0X_0/(1 + X_0) \leq 0$, i.e., $(1 + X_0)/z_0 \leq 1 - 2S_0 + 2Q_0X_0/(1 + X_0)$.

The next step is to show that $X_0 = x_0$. There are two cases here:

- $S_0 < 1/2$. Then $1 - 2S_0 + 2Q_0X_0/(1 + X_0) > 0$, so $z_0 \geq (1 + X_0)/(1 - 2S_0 + 2Q_0X_0/(1 + X_0))$. But $(1 + X_0)/(1 - 2S_0 + 2Q_0X_0/(1 + X_0)) \geq z_0$ by Theorem 3.1.2, and equality is achieved only if $X_0 = x_0$.
- $S_0 = 1/2$. Then $(1 + X_0)/z_0 \leq 2Q_0X_0/(1 + X_0)$; but $z_0 = 2Q_0/Q_0^2 = 2/Q_0$ by definition of z_0, so $(1 + X_0)^2 \leq 2Q_0X_0z_0 = 4X_0$, so $(1 - X_0)^2 \leq 0$, so $(1 - X_0)^2 = 0$, so $X_0 = 1$. Also $x_0 = Q_0/Q_0 = 1$ by definition of x_0.

These are all possible cases, since $S_0 \leq 1/2$ by hypothesis.

To recap so far, $\kappa/n \in x_0 + o(1)$ for $n \in N'''$.

A cautionary note is required at this point. If κ/n were known to have the form $x_0 + O(1)/\lg n$ then another compactness argument would extract a subsequence where κ/n has the form $x_0 + (X_1 + o(1))/\lg n$, and then finishing the proof would be a simple matter of applying Theorem 3.1.6(2). However, if one has, e.g., $\kappa/n \in x_0 + (\lg \lg n + o(1))/\lg n$ then there is no subsequence of the form $x_0 + (X_1 + o(1))/\lg n$. The rest of the proof here uses a different strategy.

Define $\rho = \text{StandardRatio}(n, q, s, \kappa, B)$. The following calculations will put a lower bound on $\lg \rho$ for $n \in N'''$ of the form $c + o(1)$ for a positive real number c. First, by definition of StandardRatio,

$$2 \lg \rho = \lg((n + \kappa)s^2 + 1) - \lg \frac{d}{B} - 2(2B - d - 1) \lg D - 2\frac{\kappa}{d} \lg q.$$

To put a lower bound on the third and fourth terms on the right side, observe that

$$\begin{aligned}(2B - d - 1) \lg D + \frac{\kappa}{d} \lg q &= (2B - n - \kappa - 2) \lg D + \frac{\kappa}{n + \kappa + 1} \lg q \\ &\leq (2B - 1) \lg D + \lg q - 2\sqrt{(n + 1)(\lg D) \lg q}\end{aligned}$$

by Theorem 3.1.3, since $2 \leq n$, $2 \leq q$, $1 \leq \kappa$, $2 \leq B$, and $1 < D$. Define

$$L = \lg((n + \kappa)s^2 + 1) - \lg \frac{d}{B} - 2(2B - 1) \lg D - 2 \lg q + 4\sqrt{(n + 1)(\lg D) \lg q};$$

then $2 \lg \rho \geq L$. The rest of the proof calculates the asymptotics of L for $n \in N'''$.

For the first and second terms, use $\kappa/n \in x_0 + o(1)$, $B/n \in z_0 + o(1)$, and $s \in S_0 \lg n + S_1 + o(1)$ to see that

$$\lg((n + \kappa)s^2 + 1) \in (1 + 2S_0) \lg n + \lg(1 + x_0) + 2S_1 + o(1)$$

and

$$\lg \frac{d}{B} \in \lg(1 + x_0) - \lg z_0 + o(1)$$

exactly as in the proof of Theorem 3.1.6, so $\lg((n + \kappa)s^2 + 1) - \lg(d/B) \in (1 + 2S_0) \lg n + \lg z_0 + 2S_1 + o(1)$.

For the $2(2B - 1) \lg D$ term, note that $(2B - 1)/(B - 1) \in 2 + o(1)/\lg n$ and that

$$2(B - 1) \lg D = \lg B + \frac{\lg \pi B}{B} - \lg(2\pi \exp 1) \in \lg n + \lg z_0 - \lg(2\pi \exp 1) + o(1)$$

so $2(2B-1)\lg D \in 2\lg n + 2(\lg z_0 - \lg(2\pi\exp 1)) + o(1)$.

For the $\sqrt{(n+1)(\lg D)\lg q}$ term, first use

$$\frac{B-1}{n+1} \in z_0 + \frac{z_1 - \epsilon + o(1)}{\lg n}$$

to see that $(n+1)/(B-1) \in 1/z_0 + ((\epsilon - z_1)/z_0^2 + o(1))/\lg n$. Multiply by z_0 and by

$$2(B-1)\lg D \in \lg n + \lg \frac{z_0}{2\pi\exp 1} + o(1)$$

to see that

$$2z_0(n+1)\lg D \in \lg n + \frac{\epsilon}{z_0} - \frac{z_1}{z_0} + \lg \frac{z_0}{2\pi\exp 1} + o(1).$$

By definition $z_0 = 2Q_0/(Q_0 - S_0 + 1/2)^2$ and

$$\begin{aligned} z_1 &= \left(2S_1 + \lg z_0 - \left(S_0 - Q_0 + \frac{3}{2}\right)\lg \frac{z_0}{2\pi\exp 1} - \frac{Q_1(Q_0 + S_0 - \frac{1}{2})}{Q_0}\right)\frac{2Q_0}{(Q_0 - S_0 + \frac{1}{2})^3} \\ &= \left(2S_1 + \lg z_0 - \left(S_0 - Q_0 + \frac{3}{2}\right)\lg \frac{z_0}{2\pi\exp 1} - \frac{Q_1(Q_0 + S_0 - \frac{1}{2})}{Q_0}\right)\frac{z_0}{Q_0 - S_0 + \frac{1}{2}} \end{aligned}$$

so

$$\begin{aligned} \frac{z_1}{z_0} - \lg \frac{z_0}{2\pi\exp 1} &= \left(2S_1 + \lg z_0 - 2\lg \frac{z_0}{2\pi\exp 1} - \frac{Q_1(Q_0 + S_0 - \frac{1}{2})}{Q_0}\right)\frac{1}{Q_0 - S_0 + \frac{1}{2}} \\ &= \left(2S_1 + \lg z_0 - 2\lg \frac{z_0}{2\pi\exp 1}\right)\frac{1}{Q_0 - S_0 + \frac{1}{2}} - \frac{Q_1 x_0}{Q_0} \\ &= \left(2S_1 + \lg z_0 - 2\lg \frac{z_0}{2\pi\exp 1}\right)\frac{1+x_0}{2Q_0} - \frac{Q_1 x_0}{Q_0}. \end{aligned}$$

Hence

$$2z_0(n+1)\lg D \in \lg n + \frac{\epsilon}{z_0} - \left(2S_1 + \lg z_0 - 2\lg \frac{z_0}{2\pi\exp 1}\right)\frac{1+x_0}{2Q_0} + \frac{Q_1 x_0}{Q_0} + o(1).$$

Multiply by $(\lg q)/Q_0 \in \lg n + Q_1/Q_0 + o(1)$:

$$\begin{aligned} &\frac{2z_0}{Q_0}(n+1)(\lg D)\lg q \in (\lg n)^2 \\ &\quad + \left(\frac{Q_1}{Q_0} + \frac{\epsilon}{z_0} - \left(2S_1 + \lg z_0 - 2\lg \frac{z_0}{2\pi\exp 1}\right)\frac{1+x_0}{2Q_0} + \frac{Q_1 x_0}{Q_0} + o(1)\right)\lg n \end{aligned}$$

Rewrite $2z_0/Q_0$ as $(1+x_0)^2/Q_0^2$, merge the Q_1 terms, and take square roots:

$$\begin{aligned} &\frac{1+x_0}{Q_0}\sqrt{(n+1)(\lg D)\lg q} \\ &\quad \in \lg n + \frac{(1+x_0)Q_1}{2Q_0} + \frac{\epsilon}{2z_0} - \left(2S_1 + \lg z_0 - 2\lg \frac{z_0}{2\pi\exp 1}\right)\frac{1+x_0}{4Q_0} + o(1). \end{aligned}$$

Multiply by $4Q_0/(1+x_0)$:

$$\begin{aligned} &4\sqrt{(n+1)(\lg D)\lg q} \\ &\quad \in \frac{4Q_0}{1+x_0}\lg n + 2Q_1 + \frac{2Q_0\epsilon}{z_0(1+x_0)} - 2S_1 - \lg z_0 + 2\lg \frac{z_0}{2\pi\exp 1} + o(1). \end{aligned}$$

Now add everything:

$$\begin{aligned}L \in {} & (1+2S_0)\lg n + \lg z_0 + 2S_1 + o(1) \\ & - 2\lg n - 2\lg\frac{z_0}{2\pi\exp 1} + o(1) \\ & - 2Q_0\lg n - 2Q_1 + o(1) \\ & + \frac{4Q_0}{1+x_0}\lg n + 2Q_1 + \frac{2Q_0\epsilon}{z_0(1+x_0)} - 2S_1 - \lg z_0 + 2\lg\frac{z_0}{2\pi\exp 1} + o(1).\end{aligned}$$

The coefficients of $\lg n$ cancel, since $1+2S_0-2-2Q_0 = -2(Q_0-S_0+1/2) = -4Q_0/(1+x_0)$. Almost all of the remaining terms also cancel, leaving just

$$L \in \frac{2Q_0\epsilon}{z_0(1+x_0)} + o(1).$$

All of $Q_0, \epsilon, z_0, 1+x_0$ are positive, so $L > 0$ for all sufficiently large n, so $\rho > 1$ for all sufficiently large n; but by assumption $\rho \le 1$ for all sufficiently large n. Contradiction.

Consequently, for each $\epsilon > 0$, there are only finitely many $n \in N$ for which $\beta/n \le z_0 + (z_1 - \epsilon)/\lg n$. By Theorem 3.1.1, there is a function $n \mapsto \ell$ with $\beta \ge \ell$ and $\ell/n \in z_0 + (z_1 + o(1))/\lg n$ as claimed. □

References

[A1C18] Carlisle Adams, Jan Camenisch (editors), *Selected areas in cryptography—SAC 2017, 24th international conference, Ottawa, ON, Canada, August 16–18, 2017, revised selected papers*, Lecture Notes in Computer Science, 10719, Springer, 2018. ISBN 978-3-319-72564-2. See [B4CLV18].

[A2.91] Leonard M. Adleman, *Factoring numbers using singular integers*, in STOC 1991 [**K5V91**] (1991), 64–71. DOI: 10.1145/103418.103432. Citations in this document: §1, §1.

[A3.17] Martin R. Albrecht, *On dual lattice attacks against small-secret LWE and parameter choices in HElib and SEAL*, in Eurocrypt 2017 [**C10N17**] (2017), 103–129. URL: https://eprint.iacr.org/2017/047. DOI: 10.1007/978-3-319-56614-6_4. Citations in this document: §1.4, §1.4.

[A3CDD+18] Martin R. Albrecht, Benjamin R. Curtis, Amit Deo, Alex Davidson, Rachel Player, Eamonn W. Postlethwaite, Fernando Virdia, Thomas Wunderer, *Estimate all the {LWE, NTRU} Schemes!*, in SCN 2018 [**C3P18**] (2018), 351–367. URL: https://eprint.iacr.org/2018/331. DOI: 10.1007/978-3-319-98113-0_19. Citations in this document: §2, §2.6, §2.6, §2.6.

[A3D21] Martin R. Albrecht, Léo Ducas, *Lattice attacks on NTRU and LWE: a history of refinements* (2021). URL: https://eprint.iacr.org/2021/799. Citations in this document: §1.1, §A.2.

[A3GVW17] Martin R. Albrecht, Florian Göpfert, Fernando Virdia, Thomas Wunderer, *Revisiting the expected cost of solving uSVP and applications to LWE*, in Asiacrypt 2017 [**T1P17**] (2017), 297–322. URL: https://eprint.iacr.org/2017/815. DOI: 10.1007/978-3-319-70694-8_11. Citations in this document: §2.3.

[A4BDL+21] Erdem Alkim, Joppe W. Bos, Léo Ducas, Patrick Longa, Ilya Mironov, Michael Naehrig, Valeria Nikolaenko, Chris Peikert, Ananth Raghunathan, Douglas Stebila, *FrodoKEM: Learning With Errors key encapsulation: algorithm specifications and supporting documentation* (2021). URL: https://web.archive.org/web/20220119174856/https://frodokem.org/files/FrodoKEM-specification-20210604.pdf. Citations in this document: §1.1.

[A4DPS16] Erdem Alkim, Léo Ducas, Thomas Pöppelmann, Peter Schwabe, *Post-quantum key exchange—a new hope*, in USENIX 2016 [**H5S16**] (2016), 327–343, 9 August 2016 version. URL: https://eprint.iacr.org/2015/1092. Citations in this document: §1.1, §1.1, §1.2, §1.2, §1.4.

[A5BDK+20] Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrede Lepoint, Vadim Lyubashevsky, John M. Schanck, Peter Schwabe, Gregor Seiler, Damien Stehlé, *CRYSTALS-Kyber: Algorithm specifications and supporting documentation* (2020). URL: https://web.archive.org/web/20211007045636/https://pq-crystals.org/kyber/data/kyber-specification-round3.pdf. Citations in this document: §1.4.

[B1GJE14] Razvan Barbulescu, Pierrick Gaudry, Antoine Joux, Emmanuel Thomé, *A heuristic quasi-polynomial algorithm for discrete logarithm in finite fields of small characteristic*, in Eurocrypt 2014 [**N1O14**] (2014), 1–16. URL: https://www.iacr.org/archive/eurocrypt2014/84410132/84410132.pdf. DOI: 10.1007/978-3-642-55220-5_1. Citations in this document: §1.

[B2DGL16] Anja Becker, Léo Ducas, Nicolas Gama, Thijs Laarhoven, *New directions in nearest neighbor searching with applications to lattice sieving*, in SODA 2016 [**K6.16**] (2016), 10–24. URL: https://eprint.iacr.org/2015/1128. DOI: 10.1137/1.9781611974331.CH2. Citations in this document: §1.1.

[B2GJ14] Anja Becker, Nicolas Gama, Antoine Joux, *A sieve algorithm based on overlattices*, LMS Journal of Computation and Mathematics **17** (2014), 49–70. URL: https://eprint.iacr.org/2013/685. DOI: 10.1112/s1461157014000229. Citations in this document: §1.1.

[B3DG20] Mihir Bellare, Hannah Davis, Felix Günther, *Separate your domains: NIST PQC KEMs, oracle cloning and read-only indifferentiability*, in Eurocrypt 2020 [**C1I20**] (2020), 3–32. URL: https://eprint.iacr.org/2020/241. DOI: 10.1007/978-3-030-45724-2_1. Citations in this document: §2.6.

[B4.23] Daniel J. Bernstein, *Multi-ciphertext security degradation for lattices* (2023). URL: https://cr.yp.to/papers.html#lprrr. Citations in this document: §1.5.

[B4CLV18] Daniel J. Bernstein, Chitchanok Chuengsatiansup, Tanja Lange, Christine van Vredendaal, *NTRU Prime: reducing attack surface at low cost*, in SAC 2017 [**A1C18**] (2018), 235–260. URL: https://ntruprime.cr.yp.to/papers.html. DOI: 10.1007/978-3-319-72565-9_12. Citations in this document: §1.4, §2.1.

[B4L14] Daniel J. Bernstein, Tanja Lange, *Batch NFS*, in SAC 2014 [**JY14**] (2014), 38–58. URL: https://cr.yp.to/papers.html#batchnfs. DOI: 10.1007/978-3-319-13051-4_3. Citations in this document: §1.

[B4LS11] Daniel J. Bernstein, Tanja Lange, Peter Schwabe, *On the correct use of the negation map in the Pollard rho method*, in PKC 2011 [**C3FGN11**] (2011), 128–146. URL: https://cr.yp.to/papers.html#negation. DOI: 10.1007/978-3-642-19379-8_8. Citations in this document: §1.

[B5GGH+20] Fabrice Boudot, Pierrick Gaudry, Aurore Guillevic, Nadia Heninger, Emmanuel Thomé, Paul Zimmermann, *Comparing the difficulty of factorization and discrete logarithm: A 240-digit experiment*, in Crypto 2020 [**M5R20**] (2020), 62–91. URL: https://eprint.iacr.org/2020/697. DOI: 10.1007/978-3-030-56880-1_3. Citations in this document: §1.

[B6.98] Joe P. Buhler (editor), *Algorithmic number theory, third international symposium, ANTS-III, Portland, Oregon, USA, June 21–25, 1998, proceedings*, Lecture Notes in Computer Science, 1423, Springer, 1998. ISBN 3-540-64657-4. See [H4PS1998].

[B6LP93] Joe P. Buhler, Hendrik W. Lenstra, Jr., Carl Pomerance, *Factoring integers with the number field sieve*, in [**L4L93**] (1993), 50–94. URL: https://www.math.leidenuniv.nl/~hwl/PUBLICATIONS/1993e/art.pdf. DOI: 10.1007/BFb0091539. Citations in this document: §1.

[C1I20] Anne Canteaut, Yuval Ishai (editors), *Advances in cryptology—EUROCRYPT 2020—39th annual international conference on the theory and applications of cryptographic techniques, Zagreb, Croatia, May 10–14, 2020, proceedings, part II*, Lecture Notes in Computer Science, 12106, Springer, 2020. ISBN 978-3-030-45723-5. See [B3DG20].

[C2MR24] Claude Carlet, Kalikinkar Mandal, Vincent Rijmen (editors), *Selected areas in cryptography—SAC 2023—30th international conference, Fredericton, Canada, August 14–18, 2023, revised selected papers*, 14201, Springer, 2024. ISBN 978-3-031-53367-9. DOI: 10.1007/978-3-031-53368-6. See [C6CHY24].

[C3FGN11] Dario Catalano, Nelly Fazio, Rosario Gennaro, Antonio Nicolosi (editors), *Public key cryptography—PKC 2011—14th international conference on practice and theory in public key cryptography, Taormina, Italy, March 6-9, 2011, proceedings*, Lecture Notes in Computer Science, 6571, Springer, 2011. See [B4LS11].

[C3P18] Dario Catalano, Roberto De Prisco (editors), *Security and cryptography for networks—11th international conference, SCN 2018, Amalfi, Italy, September 5–7, 2018, proceedings*, Lecture Notes in Computer Science, 11035, Springer, 2018. See [A3CDD+18].

[C4DLL+00] Stefania Cavallar, Bruce Dodson, Arjen K. Lenstra, Walter M. Lioen, Peter L. Montgomery, Brian Murphy, Herman te Riele, Karen Aardal, Jeff Gilchrist, Gérard Guillerm, Paul C. Leyland, Joël Marchand, Francois Morain, Alec Muffett, Chris Putnam, Craig Putnam, Paul Zimmermann, *Factorization of a 512-bit RSA modulus*, in Eurocrypt 2000 [**P2.00**] (2000), 1–18. URL: https://www.iacr.org/archive/eurocrypt2000/1807/18070001-new.pdf. DOI: 10.1007/3-540-45539-6_1. Citations in this document: §1, §1.

[C5L21] André Chailloux, Johanna Loyer, *Lattice sieving via quantum random walks*, in Asiacrypt 2021 [**T2W21**] (2021), 63–91. URL: https://eprint.

iacr.org/2021/570. DOI: 10.1007/978-3-030-92068-5_3. Citations in this document: §1.1.

[C6CHY24] Jung Hee Cheon, Hyeongmin Choe, Dongyeon Hong, MinJune Yi, *SMAUG: pushing lattice-based key encapsulation mechanisms to the limits*, in SAC 2023 [**C2MR24**] (2024), 127–146. URL: https://eprint.iacr.org/2023/739. DOI: 10.1007/978-3-031-53368-6_7. Citations in this document: §1.1.

[C7HSW11] Bruce S. N. Cheung, Lucas Chi Kwong Hui, Ravi S. Sandhu, Duncan S. Wong (editors), *Proceedings of the 6th ACM symposium on information, computer and communications security, ASIACCS 2011, Hong Kong, China, March 22–24, 2011*, ACM, 2011. ISBN 978-1-4503-0564-8. See [WLTB11].

[C8.84] Don Coppersmith, *Fast evaluation of logarithms in fields of characteristic two*, IEEE Transactions on Information Theory **30** (1984), 587–594. MR 85h:65041. URL: https://pages.cs.wisc.edu/~cs812-1/coppersmith.pdf. DOI: 10.1109/TIT.1984.1056941. Citations in this document: §1.

[C8.93] Don Coppersmith, *Modifications to the number field sieve*, Journal of Cryptology **6** (1993), 169–180. URL: https://link.springer.com/content/pdf/10.1007/bf00198464.pdf. DOI: 10.1007/BF00198464. Citations in this document: §1.

[C8S97] Don Coppersmith, Adi Shamir, *Lattice attacks on NTRU*, in Eurocrypt 1997 [**F97**] (1997), 52–61. URL: https://link.springer.com/content/pdf/10.1007/3-540-69053-0_5.pdf. DOI: 10.1007/3-540-69053-0_5. Citations in this document: §2.2.

[C9LRS09] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, *Introduction to algorithms*, 3rd edition, MIT Press, 2009. ISBN 978-0262033848. Citations in this document: §1, §1.

[C10N17] Jean-Sébastien Coron, Jesper Buus Nielsen (editors), *Advances in cryptology—EUROCRYPT 2017—36th annual international conference on the theory and applications of cryptographic techniques, Paris, France, April 30–May 4, 2017, proceedings, part II*, 10211, 2017. ISBN 978-3-319-56613-9. See [A3.17].

[D1S22] Yevgeniy Dodis, Thomas Shrimpton (editors), *Advances in cryptology—CRYPTO 2022—42nd annual international cryptology conference, CRYPTO 2022, Santa Barbara, CA, USA, August 15–18, 2022, proceedings, part II*, 13508, Springer, 2022. ISBN 978-3-031-15978-7. DOI: 10.1007/978-3-031-15979-4. See [ETWY22].

[D2LW20] Emmanouil Doulgerakis, Thijs Laarhoven, Benne de Weger, *Sieve, enumerate, slice, and lift: hybrid lattice algorithms for SVP via CVPP*, in Africacrypt 2020 [**N3Y20**] (2020), 301–320. URL: https://eprint.iacr.org/2020/487. DOI: 10.1007/978-3-030-51938-4_15. Citations in this document: §1.4.

[D3.18] Léo Ducas, *Shortest vector from lattice sieving: A few dimensions for free*, in Eurocrypt 2018 [**N2R18**] (2018), 125–145. URL: https://eprint.iacr.org/2017/999. DOI: 10.1007/978-3-319-78381-9_5. Citations in this document: §1.4.

[D3P23] Léo Ducas, Ludo N. Pulles, *Does the dual-sieve attack on learning with errors even work?*, in Crypto 2023 [**H1L23**] (2023), 37–69. URL: https://eprint.iacr.org/2023/302. DOI: 10.1007/978-3-031-38548-3_2. Citations in this document: §1.4, §1.4.

[D4HKLS21] Julien Duman, Kathrin Hövelmanns, Eike Kiltz, Vadim Lyubashevsky, Gregor Seiler, *Faster lattice-based KEMs via a generic Fujisaki-Okamoto transform using prefix hashing*, in CCS 2021 [**K2KVS21**] (2021), 2722–2737. URL: https://eprint.iacr.org/2021/1351. DOI: 10.1145/3460120.3484819. Citations in this document: §2.6.

[ETWY22] Thomas Espitau, Mehdi Tibouchi, Alexandre Wallet, Yang Yu, *Shorter hash-and-sign lattice-based signatures*, in Crypto 2022 [**D1S22**] (2022), 245–275. URL: https://eprint.iacr.org/2022/785. DOI: 10.1007/978-3-031-15979-4_9. Citations in this document: §1.1.

[F97] Walter Fumy (editor), *Advances in cryptology—EUROCRYPT '97, international conference on the theory and application of cryptographic techniques, Konstanz, Germany, May 11–15, 1997*, Lecture Notes in Computer Science, 1233, Springer, 1997. See [C8S97].

[G1R15] Rosario Gennaro, Matthew Robshaw (editors), *Advances in cryptology—CRYPTO 2015—35th annual cryptology conference, Santa Barbara, CA, USA, August 16–20, 2015, proceedings, part I*, 9215, Springer, 2015. ISBN 978-3-662-47988-9. DOI: 10.1007/978-3-662-47989-6. See [L1.15].

[G2KZ18] Robert Granger, Thorsten Kleinjung, Jens Zumbrägel, *On the discrete logarithm problem in finite fields of fixed characteristic*, Transactions of the American Mathematical Society **370** (2018), 3129–3145. URL: https://eprint.iacr.org/2015/685. DOI: 10.1090/tran/7027. Citations in this document: §1.

[G3J21] Qian Guo, Thomas Johansson, *Faster dual lattice attacks for solving LWE with applications to CRYSTALS*, in Asiacrypt 2021 [**T2W21**] (2021), 33–62. DOI: 10.1007/978-3-030-92068-5_2. Citations in this document: §1.4.

[H1L23] Helena Handschuh, Anna Lysyanskaya (editors), *Advances in cryptology—CRYPTO 2023—43rd annual international cryptology conference, CRYPTO 2023, Santa Barbara, CA, USA, August 20–24, 2023, proceedings, part III*, 14083, Springer, 2023. ISBN 978-3-031-38547-6. DOI: 10.1007/978-3-031-38548-3. See [D3P23].

[H2.96] John Harrison, *HOL Light: A tutorial introduction*, in FMCAD 1996 [**S3C96**] (1996), 265–269. URL: https://www.cl.cam.ac.uk/~jrh13/papers/demo.pdf. DOI: 10.1007/BFB0031814. Citations in this document: §1.2.

[H3.21] Max Heiser, *Improved quantum hypercone locality sensitive filtering in lattice sieving* (2021). URL: https://eprint.iacr.org/2021/1295. Citations in this document: §1.1.

[H4PS1998] Jeffrey Hoffstein, Jill Pipher, Joseph H. Silverman, *NTRU: a ring-based public key cryptosystem*, in ANTS 1998 [**B6.98**] (1998), 267–288. URL: https://ntru.org/f/hps98.pdf. DOI: 10.1007/BFB0054868. Citations in this document: §2.2, §2.2, §2.2.

[H4PS2016] Jeffrey Hoffstein, Jill Pipher, Joseph H. Silverman, *NTRU: a new high speed public key cryptosystem* (2016), Circulated at Crypto 1996, put online in 2016. URL: https://ntru.org/f/hps96.pdf. Citations in this document: §2.2.

[H5S16] Thorsten Holz, Stefan Savage (editors), *25th USENIX Security Symposium, USENIX Security 16, Austin, TX, USA, August 10–12, 2016*, USENIX Assocation, 2016. See [A4DPS16].

[J14] Antoine Joux, *A new index calculus algorithm with complexity* $L(1/4 + o(1))$ *in small characteristic*, in SAC 2013 [**L2LL14**] (2014), 355–379. URL: https://eprint.iacr.org/2013/095. DOI: 10.1007/978-3-662-43414-7_18. Citations in this document: §1.

[JY14] Antoine Joux, Amr M. Youssef (editors), *Selected areas in cryptography—SAC 2014—21st international conference, Montreal, QC, Canada, August 14–15, 2014, revised selected papers*, 8781, Springer, 2014. ISBN 978-3-319-13050-7. See [B4L14].

[K1.11] Aggelos Kiayias (editor), *Topics in cryptology—CT-RSA 2011—the cryptographers' track at the RSA Conference 2011, San Francisco, CA, USA, February 14–18, 2011, proceedings*, Lecture Notes in Computer Science, 6558, Springer, 2011. ISBN 978-3-642-19073-5. See [L6P11].

[K2KVS21] Yongdae Kim, Jong Kim, Giovanni Vigna, Elaine Shi (editors), *CCS '21: 2021 ACM SIGSAC conference on computer and communications security, virtual event, Republic of Korea, November 15–19, 2021*, ACM, 2021. ISBN 978-1-4503-8454-4. See [D4HKLS21].

[K3L21] Elena Kirshanova, Thijs Laarhoven, *Lower bounds on lattice sieving and information set decoding*, in Crypto 2021 [**M1P21**] (2021), 791–820. URL: https://eprint.iacr.org/2021/785. DOI: 10.1007/978-3-030-84245-1_27. Citations in this document: §1.1.

[K4AFL+10] Thorsten Kleinjung, Kazumaro Aoki, Jens Franke, Arjen K. Lenstra, Emmanuel Thomé, Joppe W. Bos, Pierrick Gaudry, Alexander Kruppa, Peter L. Montgomery, Dag Arne Osvik, Herman J. J. te Riele, Andrey Timofeev, Paul Zimmermann, *Factorization of a 768-bit RSA modulus*, in Crypto 2010 [**R10**] (2010), 333–350. URL: https://eprint.iacr.org/2010/006. DOI: 10.1007/978-3-642-14623-7_18. Citations in this document: §1.

[K5V91] Cris Koutsougeras, Jeffrey Scott Vitter (editors), *Proceedings of the 23rd annual ACM symposium on theory of computing, May 5–8, 1991, New Orleans, Louisiana, USA*, Association for Computing Machinery, New York, 1991. ISBN 0-89791-397-3. See [A2.91].

[K6.16] Robert Krauthgamer (editor), *Proceedings of the twenty-seventh annual ACM-SIAM symposium on discrete algorithms, SODA 2016, Arlington, VA, USA, January 10–12, 2016*, SIAM, 2016. ISBN 978-1-61197-433-1. See [B2DGL16].

[K7S01] Fabian Kuhn, Rene Struik, *Random walks revisited: extensions of Pollard's rho algorithm for computing multiple discrete logarithms*, in SAC 2001 [**VY01**] (2001), 212–229. URL: http://www.distcomp.ethz.ch/publications.html. DOI: 10.1007/3-540-45537-X_17. Citations in this document: §1.

[L1.15] Thijs Laarhoven, *Sieving for shortest vectors in lattices using angular locality-sensitive hashing*, in Crypto 2015 [**G1R15**] (2015), 3–22. URL: https://eprint.iacr.org/2014/744. DOI: 10.1007/978-3-662-47989-6_1. Citations in this document: §1.1.

[L1W15] Thijs Laarhoven, Benne de Weger, *Faster sieving for shortest lattice vectors using spherical locality-sensitive hashing*, in LATINCRYPT 2015 [**L3R15**] (2015), 101–118. URL: https://eprint.iacr.org/2015/211. DOI: 10.1007/978-3-319-22174-8_6. Citations in this document: §1.1.

[L2LL14] Tanja Lange, Kristin E. Lauter, Petr Lisonek (editors), *Selected areas in cryptography—SAC 2013—20th international conference, Burnaby, BC, Canada, August 14–16, 2013, revised selected papers*, 8282, Springer, 2014. ISBN 978-3-662-43413-0. DOI: 10.1007/978-3-662-43414-7. See [J14], [ZPH14].

[L3R15] Kristin E. Lauter, Francisco Rodríguez-Henríquez (editors), *Progress in cryptology—LATINCRYPT 2015—4th international conference on cryptology and information security in Latin America, Guadalajara, Mexico, August 23–26, 2015, proceedings*, 9230, Springer, 2015. ISBN 978-3-319-22173-1. DOI: 10.1007/978-3-319-22174-8. See [L1W15].

[L4L93] Arjen K. Lenstra, Hendrik W. Lenstra, Jr. (editors), *The development of the number field sieve*, Lecture Notes in Mathematics, 1554, Springer, 1993. ISBN 3-540-57013-6. MR 96m:11116. DOI: 10.1007/BFb0091534. Citations in this document: §1. See [B6LP93].

[L4LMP93] Arjen K. Lenstra, Hendrik W. Lenstra, Jr., Mark S. Manasse, John M. Pollard, *The number field sieve*, 1993. URL: https://www.math.leidenuniv.nl/~hwl/PUBLICATIONS/1993d/art.pdf. DOI: 10.1007/BFb0091537. Citations in this document: §1.

[L5P92] Hendrik W. Lenstra, Jr., Carl Pomerance, *A rigorous time bound for factoring integers*, Journal of the American Mathematical Society **5** (1992), 483–516. URL: https://www.ams.org/journals/jams/1992-05-03/S0894-0347-1992-1137100-0/. DOI: 10.1090/S0894-0347-1992-1137100-0. Citations in this document: §1.

[L6P11] Richard Lindner, Chris Peikert, *Better key sizes (and attacks) for LWE-based encryption*, in CT-RSA [**K1.11**] (2011), 319–339. URL: https://eprint.iacr.org/2010/613. DOI: 10.1007/978-3-642-19074-2_21. Citations in this document: §2.5.

[L7PR13] Vadim Lyubashevsky, Chris Peikert, Oded Regev, *On ideal lattices and learning with errors over rings*, Journal of the ACM **60** (2013), Article 43, 35 pages. URL: https://eprint.iacr.org/2012/230. DOI: 10.1145/2535925. Citations in this document: §1.1, §2.1, §2.1, §2.1, §2.1, §2.2, §2.2.

[M1P21] Tal Malkin, Chris Peikert (editors), *Advances in cryptology—CRYPTO 2021—41st annual international cryptology conference, CRYPTO 2021, virtual event, August 16–20, 2021, proceedings, part II*, 12826, Springer, 2021. ISBN 978-3-030-84244-4. See [K3L21].

[M2.09] Mitsuru Matsui (editor), *Advances in cryptology—ASIACRYPT 2009, 15th international conference on the theory and application of cryptology and*

information security, Tokyo, Japan, December 6–10, 2009. proceedings, 5912, Springer, 2009. ISBN 978-3-642-10365-0. See [S4STX09].

[M3.22] MATZOV, *Report on the security of LWE: improved dual lattice attack* (2022). DOI: 10.5281/zenodo.6412487. Citations in this document: §1.4.

[M4S01] Alexander May, Joseph H. Silverman, *Dimension reduction methods for convolution modular lattices*, in CaLC 2001 [**S1.01**] (2001), 110–15. URL: https://www.cits.ruhr-uni-bochum.de/personen/may/publications.html. DOI: 10.1007/3-540-44670-2_10. Citations in this document: §2.2.

[M5R20] Daniele Micciancio, Thomas Ristenpart (editors), *Advances in cryptology—CRYPTO 2020—40th annual international cryptology conference, CRYPTO 2020, Santa Barbara, CA, USA, August 17–21, 2020, proceedings, part II*, 12171, Springer, 2020. ISBN 978-3-030-56879-5. DOI: 10.1007/978-3-030-56880-1. See [B5GGH+20].

[N1O14] Phong Q. Nguyen, Elisabeth Oswald (editors), *Advances in cryptology—EUROCRYPT 2014—33rd annual international conference on the theory and applications of cryptographic techniques, Copenhagen, Denmark, May 11–15, 2014. proceedings*, 8441, Springer, 2014. ISBN 978-3-642-55219-9. DOI: 10.1007/978-3-642-55220-5. See [B1GJE14].

[N1V08] Phong Q. Nguyen, Thomas Vidick, *Sieve algorithms for the shortest vector problem are practical*, Journal of Mathematical Cryptology **2** (2008), 181–207. URL: https://people.csail.mit.edu/vidick/JoMC08.pdf. DOI: 10.1515/JMC.2008.009. Citations in this document: §1.1.

[N2R18] Jesper Buus Nielsen, Vincent Rijmen (editors), *Advances in cryptology—EUROCRYPT 2018—37th annual international conference on the theory and applications of cryptographic techniques, Tel Aviv, Israel, April 29–May 3, 2018 proceedings, part I*, 10820, Springer, 2018. ISBN 978-3-319-78380-2. See [D3.18].

[N3Y20] Abderrahmane Nitaj, Amr M. Youssef (editors), *Progress in cryptology—AFRICACRYPT 2020—12th international conference on cryptology in Africa, Cairo, Egypt, July 20–22, 2020, proceedings*, 12174, Springer, 2020. ISBN 978-3-030-51937-7. See [D2LW20].

[OW99] Paul C. van Oorschot, Michael Wiener, *Parallel collision search with cryptanalytic applications*, Journal of Cryptology **12** (1999), 1–28. ISSN 0933–2790. DOI: 10.1007/PL00003816. Citations in this document: §1.

[P1S23] Amaury Pouly, Yixin Shen, *Provable dual attacks on learning with errors* (2023). URL: https://eprint.iacr.org/2023/1508. Citations in this document: §1.4.

[P2.00] Bart Preneel (editor), *Advances in cryptology—EUROCRYPT 2000, international conference on the theory and application of cryptographic techniques, Bruges, Belgium, May 14–18, 2000*, Lecture Notes in Computer Science, 1807, Springer, 2000. ISBN 3-540-67517-5. See [C4DLL+00].

[R10] Tal Rabin (editor), *Advances in cryptology—CRYPTO 2010, 30th annual cryptology conference, Santa Barbara, CA, USA, August 15–19, 2010, proceedings*, 6223, Springer, 2010. ISBN 978-3-642-14622-0. DOI: 10.1007/978-3-642-14623-7. See [K4AFL+10].

[S1.01] Joseph H. Silverman (editor), *Cryptography and lattices: proceedings of the 1st International Conference (CaLC 2001) held in Providence, RI, March 29–30, 2001*, Lecture Notes in Computer Science, 2146, Springer, 2001. ISBN 3-540-42488-1. MR 2002m:11002. See [M4S01].

[S2.20] Steven S. Skiena, *The algorithm design manual*, 3rd edition, Springer, 2020. ISBN 978-3-030-54255-9. Citations in this document: §1.

[S3C96] Mandayam K. Srivas, Albert John Camilleri (editors), *Formal methods in computer-aided design, first international conference, FMCAD '96, Palo Alto, California, USA, November 6–8, 1996, proceedings*, Lecture Notes in Computer Science, 1166, Springer, 1996. ISBN 3-540-61937-2. See [H2.96].

[S4STX09] Damien Stehlé, Ron Steinfeld, Keisuke Tanaka, Keita Xagawa, *Efficient public key encryption based on ideal lattices*, in Asiacrypt 2009 [**M2.09**] (2009), 617–635. URL: https://eprint.iacr.org/2009/285. DOI: 10.1007/978-3-642-10366-7_36. Citations in this document: §2.2, §2.2.

[T1P17] Tsuyoshi Takagi, Thomas Peyrin (editors), *Advances in cryptology—ASIACRYPT 2017—23rd international conference on the theory and applications of cryptology and information security, Hong Kong, China, December 3–7, 2017, proceedings, part II*, Lecture Notes in Computer Science, 10625, Springer, 2017. ISBN 978-3-319-70696-2. See [A3GVW17].

[T2W21] Mehdi Tibouchi, Huaxiong Wang (editors), *Advances in cryptology—ASIACRYPT 2021—27th international conference on the theory and application of cryptology and information security, Singapore, December 6–10, 2021, proceedings, part IV*, 13093, Springer, 2021. ISBN 978-3-030-92067-8. See [C5L21], [G3J21].

[VY01] Serge Vaudenay, Amr M. Youssef (editors), *Selected areas in cryptography: 8th annual international workshop, SAC 2001, Toronto, Ontario, Canada, August 16–17, 2001, revised papers*, Lecture Notes in Computer Science, 2259, Springer, 2001. ISBN 3-540-43066-0. MR 2004k:94066. See [K7S01].

[WLTB11] Xiaoyun Wang, Mingjie Liu, Chengliang Tian, Jingguo Bi, *Improved Nguyen-Vidick heuristic sieve algorithm for shortest vector problem*, in [**C7HSW11**] (2011), 1–9. URL: https://eprint.iacr.org/2010/647. DOI: 10.1145/1966913.1966915. Citations in this document: §1.1.

[ZPH14] Feng Zhang, Yanbin Pan, Gengran Hu, *A three-level sieve algorithm for the shortest vector problem*, in SAC 2013 [**L2LL14**] (2014), 29–47. URL: https://eprint.iacr.org/2013/536. DOI: 10.1007/978-3-662-43414-7_2. Citations in this document: §1.1.

A Computer-verified proof

As a supplement to this paper, https://cr.yp.to/2023/lprrr-20230317.ml (also attached to this PDF) presents computer-verified proofs for a generalization of Theorem 1.2.1. These proofs are written in the HOL Light language and have been verified by the HOL Light verifier.

This appendix reports how the proofs were developed, compares the computer-verified theorem statement to the statement of Theorem 1.2.1, and explains how to re-run the verifier.

A.1 Proof development

This paper is an outgrowth of the following material posted in November 2022: 2.5 pages with a theorem and proof—at a normal mathematical level of formality, not computer-verified—determining the second-order asymptotics of StandardRatio for any particular growth of (n, q, s, κ, β); and 1.5 pages informally optimizing (κ, β) and sketching how this optimization could be proven.

That initial theorem, proof, informal optimization, and optimization-proof sketch were then upgraded to a main theorem stating the asymptotics of the optimal (κ, β), and a proof of those asymptotics. The main theorem statement is Theorem 1.2.1; the proof is Section 3.

The theorem and proof were then upgraded to computer-verified proofs in the HOL Light language—along with proofs of many necessary lemmas. Here is an example of how this expanded the proofs. The proof in Section 3 includes a comment that "$1/(A_0 + (A_1 + o(1))/\lg n)$ is $1/A_0 + (-A_1/A_0^2 + o(1))/\lg n$ if $A_0 \neq 0$", where the $o(1)$ is as $n \to \infty$. Inside `lprrr-20230317.ml`, the computer-verified proof of this comment

- starts from the mean-value theorem (which is already proven in the HOL Light library);
- spends 25 lines stating and proving a more suitable two-sided version of the mean-value theorem;
- spends 74 lines stating and proving that if f is continuously differentiable at A_0 then anything in $f(A_0 + (A_1 + o(1))/X)$ is in $f(A_0) + (A_1 f'(A_0) + o(1))/X$ where the $o(1)$ is as $X \to \infty$;
- spends 39 lines specializing the statement and proof to inversion; and
- spends 17 lines specializing the statement and proof to $X = \lg n$.

Overall `lprrr-20230317.ml` occupies 345KB, nearly 10000 lines. Some of the proofs were generated by ad-hoc scripts. There are a few comments, partly for tracking the internal organization of `lprrr-20230317.ml` and partly about proofs of one of the lemmas (Bolzano's theorem). The time for writing `lprrr-20230317.ml` was an unrecorded fraction of a 3.5-week period. No claims of optimality are made for the numbers 345, 10000, and 3.5. The main theorem statement and `lprrr-20230317.ml` were posted in March 2023.

A.2 The value of computer verification

Given that the literature presents merely heuristic arguments that the standard block size is close to the actual block size required for attacks, the reader might be wondering why this paper puts so much effort into eliminating risks of error in this paper's statements about the asymptotics of the standard block size.

One answer is that the literature often describes the standard block size as an accurate approximation to the actual block size. Consider, e.g., [**A3D21**] saying that the existing heuristics were "empirically investigated and confirmed" and that various discrepancies disappear as problem sizes increase. Readers who trust the existing heuristics, on the basis of current evidence or evidence collected in the future, can—thanks to the computer-verified proofs—place the same trust in conclusions obtained by combining the heuristics with Theorem 1.2.1.

Another answer is that, even if the standard block size is somewhat inaccurate, an error buried somewhere in the asymptotic calculations in Section 3 could easily create much larger inaccuracies. The formally verified proof guarantees that Theorem 1.2.1 is correct.

A.3 The statement of the computer-verified theorem

There are two main theorems in `lprrr-20230317.ml`: `forward_main` is a generalization of Theorem 1.2.1(1), and `converse_main` is a generalization of Theorem 1.2.1(2).

A reader checking that Theorem 1.2.1 has been computer-verified must check the statements of `forward_main` and `converse_main`, along with the underlying definitions. The following paragraphs review the definitions and the theorem statements, without assuming familiarity with HOL Light.

```
let log2 = new_definition `
  log2 x = (ln x) / (ln (&2))
`;;
```

The HOL Light library already defines a function `ln`; these lines define a function `log2`. In the HOL Light language, natural numbers such as `2` are distinguished from real numbers such as `&2`. Parentheses can often be omitted but are included here for clarity.

```
let ceil = new_definition `
  ceil x = -- floor(--x)
`;;
```

This defines `ceil` on top of the function `floor` defined in the HOL Light library: $\lceil x \rceil = -\lfloor -x \rfloor$. In the HOL Light language, `--` is negation.

```
let o1_seq = new_definition `
  o1_seq (f:num->real) <=>
    !e:real. &0 < e ==>
      ?m:num.
        !i:num. m <= i ==> abs(f(i)) <= e
`;;
```

Let f be a function from $\{0, 1, 2, \ldots\}$ to $\mathbb{R}$. This definition says that `o1_seq` f is the following statement: for every $e \in \mathbb{R}$ with $0 < e$, there exists $m \in \{0, 1, 2, \ldots\}$ such that every $i \in \{0, 1, 2, \ldots\}$ with $m \leq i$ has $|f(i)| \leq e$. This is one of the traditional ways to say that f converges to 0, i.e., that $f \in o(1)$.

This is equivalent to a special case of a concept `tends_num_real` in the HOL Light library. There is only a small overlap between `o1_seq` theorems in `lprrr-20230317.ml` and `tends_num_real` theorems already in the library.

As this definition illustrates, in the HOL Light language, "`!x.`" means "for every `x` we have"; "`?x.`" means "there exists `x` such that"; "`:num`", "`:real`", and "`:num->real`" specify types. Often HOL Light can deduce types automatically, but including the types can still add clarity.

```
parse_as_infix("powreal",(24,"left"));;
let powreal = new_definition `
  x powreal y = exp(y * ln x)
`;;
```

This defines x `powreal` y as $\exp(y \ln x)$, i.e., x^y.

```
let bkzdelta = new_definition `
  bkzdelta x =
    (x * ((pi*x) powreal (&1 / x))/(&2 * pi * exp(&1)))
    powreal (&1 / (&2 * (x - &1)))
`;;
```

For comparison, $\delta = (\beta(\pi\beta)^{1/\beta}/(2\pi\exp 1))^{1/2(\beta-1)}$ inside Definition 1.1.1.

```
let standardratio = new_definition `
  standardratio n q s k x =
     ( ((n + k)*(s pow 2) + &1) powreal (&1 / &2)
     ) / ( (((n + k + &1)/x) powreal (&1 / &2))
           * ((bkzdelta x) powreal (&2 * x - (n + k + &1) - &1))
           * (q powreal (k/(n + k + &1)))
         )
`;;
```

Compare this to $\mathrm{StandardRatio}(n,q,s,\kappa,\beta) = ((n+\kappa)s^2+1)^{1/2}/(d/\beta)^{1/2}\delta^{2\beta-d-1}q^{\kappa/d}$ in Definition 1.1.1, with $d = n+\kappa+1$ and δ as above. In the HOL Light language, `pow` is exponentiation with a natural-number exponent; there is also a `sqrt(...)` that could be used in place of `(...) powreal (&1 / &2)`.

```
let forward_main = prove(`
  !n:num->real q:num->real s:num->real
   Q0:real Q1:real S0:real S1:real
   x0:real z0:real z1:real.
```

This is the start of the first main theorem statement. At a high level, the statement looks like "`!X. A /\ B ==> ?Y. C /\ D`", meaning that, for every `X` where the hypotheses `A` and `B` hold, there exists `Y` where the conclusions `C` and `D` hold. In the HOL Light language, `==>` is implication, and `/\` is conjunction.

In Theorem 1.2.1, n runs through a specified infinite subset N of $\{2,3,\ldots\}$. To match this up to the more general `n:num->real` allowed in `forward_main`, define n_0 as the smallest element of N, define n_1 as the next element of N, etc.

In Theorem 1.2.1, q and s are determined by n. The setting of `forward_main` is more general, allowing $n_i = n_j$ with $q_i \neq q_j$ or $s_i \neq s_j$.

```
   (!i. &1 < n(i))
   /\ (!i. &1 < q(i))
   /\ (!i. &0 < s(i))
   /\ o1_seq (\i. &1 / n(i))
   /\ o1_seq (\i. (log2(q(i))/log2(n(i)) - Q0) * log2(n(i)) - Q1)
   /\ o1_seq (\i. (log2(s(i))/log2(n(i)) - S0) * log2(n(i)) - S1)
```

These hypotheses say that each n_i is larger than 1; each q_i is larger than 1; each s_i is larger than 0; $1/n_i \in o(1)$ as $i \to \infty$; $((\lg q_i)/\lg n_i - Q_0)\lg n_i - Q_1 \in o(1)$, i.e., $\lg q_i \in Q_0 \lg n_i + Q_1 + o(1)$; and $\lg s_i \in S_0 \lg n_i + S_1 + o(1)$.

If n_i runs through the elements of N in order, with N as in Theorem 1.2.1, then $n_i \to \infty$ as $i \to \infty$, and $1/n_i \in o(1)$. Also, Theorem 1.2.1 assumes $\lg q \in Q_0 \lg n + Q_1 + o(1)$ and $\lg s \in S_0 \lg n + S_1 + o(1)$.

```
   /\ -- &1 / &2 < S0
   /\ S0 <= &1 / &2
   /\ &1 / &2 < Q0 - S0
   /\ &1 / &2 < Q0 + S0
```

These hypotheses say $-1/2 < S_0 \leq 1/2$, $1/2 < Q_0 - S_0$, and $1/2 < Q_0 + S_0$. All of these are satisfied in Theorem 1.2.1, which requires $0 \leq S_0 \leq 1/2 < Q_0 - S_0$.

```
/\ x0 = (Q0 + S0 - &1 / &2)/(Q0 - S0 + &1 / &2)
/\ z0 = &2 * Q0/((Q0 - S0 + &1 / &2) pow 2)
/\ z1 =   (&2 * S1 + log2(z0)
           - (S0 - Q0 + &3 / &2) * (log2(z0) - log2(&2 * pi * exp(&1)))
           - Q1 * (Q0 + S0 - &1 / &2)/Q0
          )
        * (&2 * Q0) / ((Q0 - S0 + &1 / &2) pow 3)
```

For comparison, Theorem 1.2.1 says $x_0 = (Q_0 + S_0 - 1/2)/(Q_0 - S_0 + 1/2)$; $z_0 = 2Q_0/(Q_0 - S_0 + 1/2)^2$; and

$$z_1 = \left(2S_1 + \lg z_0 - \left(S_0 - Q_0 + \frac{3}{2}\right) \lg \frac{z_0}{2\pi \exp 1} - \frac{Q_1(Q_0 + S_0 - \frac{1}{2})}{Q_0}\right) \frac{2Q_0}{(Q_0 - S_0 + \frac{1}{2})^3}.$$

The constant `pi` is provided by the HOL Light library.

```
==> ?k:num->real b:num->real.
```

This says that, if the above hypotheses are satisfied, then there exist functions k, b from $\{0, 1, 2, \ldots\}$ to $\mathbb{R}$ satisfying the conclusions that follow.

If n is an injective function on $\{0, 1, 2, \ldots\}$ then the functions $i \mapsto k_i$ and $i \mapsto b_i$ are determined by functions $n \mapsto k$ and $n \mapsto b$; Theorem 1.2.1 is phrased in terms of the latter functions.

```
  (!i. integer(k(i)))
  /\ (!i. &0 < k(i))
  /\ (!i. k(i) <= ceil(n(i)))
  /\ o1_seq (\i. (k(i)/n(i) - x0) * log2(n(i)) - &0)
```

This says that each k_i is an integer, that $0 < k_i \le \lceil n_i \rceil$, and that $k_i/n_i \in x_0 + o(1)/\lg n_i$. In particular, $1 \le k_i$, and $k_i \le n_i$ if n_i is an integer, the situation of Theorem 1.2.1.

```
  /\ (!i. integer(b(i)))
  /\ (!i. &1 < b(i))
  /\ (!i. b(i) <= ceil(n(i) + k(i) + &1))
  /\ o1_seq (\i. (b(i)/n(i) - z0) * log2(n(i)) - z1)
```

This says that each b_i is an integer, that $1 < b_i \le \lceil n_i + k_i + 1 \rceil$, and that $b_i/n_i \in z_0 + (z_1 + o(1))/\lg n_i$. In particular, $2 \le b_i$, and $b_i \le n_i + k_i + 1$ if n_i is an integer (since k_i is also an integer).

```
  /\ ?m. !i. m <= i ==>
     standardratio (n(i)) (q(i)) (s(i)) (k(i)) (b(i)) < &1
`,
...
```

This covers the last conclusion of Theorem 1.2.1(1): there is some m such that every $i \ge m$ has StandardRatio$(n_i, q_i, s_i, k_i, b_i) < 1$.

A proof in `lprrr-20230317.ml` has been replaced with `...` here. The main point of computer verification is that the reader does not need to check the proof.

```
let converse_main = prove(`
  !n:num->real q:num->real s:num->real
   k:num->real b:num->real
   Q0:real Q1:real S0:real S1:real
   x0:real z0:real z1:real.
```

This starts the other main theorem statement, generalizing Theorem 1.2.1(2). Note the extra `k` and `b` here.

```
(!i. &1 < n(i))
/\ (!i. &1 < q(i))
/\ (!i. &0 < s(i))
/\ o1_seq (\i. &1 / n(i))
/\ o1_seq (\i. (log2(q(i))/log2(n(i)) - Q0) * log2(n(i)) - Q1)
/\ o1_seq (\i. (log2(s(i))/log2(n(i)) - S0) * log2(n(i)) - S1)
```

This is exactly the same as in the first main theorem statement.

```
/\ -- &1 / &2 < S0
/\ S0 <= &1 / &2
/\ -- &1 / &2 < Q0 - S0
/\ &1 / &2 < Q0 + S0
```

This is more generous than in the first main theorem statement: this requires merely $-1/2 < Q_0 - S_0$, not $1/2 < Q_0 - S_0$.

```
/\ x0 = (Q0 + S0 - &1 / &2)/(Q0 - S0 + &1 / &2)
/\ z0 = &2 * Q0/((Q0 - S0 + &1 / &2) pow 2)
/\ z1 =   (&2 * S1 + log2(z0)
           - (S0 - Q0 + &3 / &2) * (log2(z0) - log2(&2 * pi * exp(&1)))
           - Q1 * (Q0 + S0 - &1 / &2)/Q0
          )
        * (&2 * Q0) / ((Q0 - S0 + &1 / &2) pow 3)
```

This is again exactly the same as in the first main theorem statement.

```
/\ (!i. &0 < k(i))
/\ (!i. k(i) <= &100 * n(i))
/\ (!i. &60 <= b(i))
/\ (!i. b(i) <= ceil(n(i) + k(i) + &1))
/\ (?m. !i. m <= i ==>
    standardratio (n(i)) (q(i)) (s(i)) (k(i)) (b(i)) <= &1)
```

This says $0 < k_i \le 100n_i$ and $60 \le b_i \le \lceil n_i + k_i + 1 \rceil$. These inequalities are satisfied if $1 \le k_i \le 100n_i$ and $60 \le b_i \le n_i + k_i + 1$, as in Theorem 1.2.1(2).

This also says that, for all suficiently large i, $\mathrm{StandardRatio}(n_i, q_i, s_i, k_i, b_i) \le 1$. This is assumed by Theorem 1.2.1(2).

```
==>
?L. (!i. L(i) <= b(i))
/\ o1_seq (\i. (L(i)/n(i) - z0) * log2(n(i)) - z1)
`,
...
```

For comparison, the conclusion of Theorem 1.2.1(2) is that $\beta \ge \ell$ for some function $n \mapsto \ell$ with $\ell/n \in z_0 + (z_1 + o(1))/\lg n$.

A.4 Redoing the computer verification

Readers are cautioned that, beyond the portion of HOL Light responsible for verifying theorems, there are many more lines of code in the HOL Light library providing proof tools and specific proofs—and perhaps doing something else, since all of this is written in a general-purpose programming language. Malicious code in HOL Light or in this paper's `lprrr-20230317.ml` could exfiltrate secret files, install ransomware, or, perhaps most terrifyingly, output a "`thm`" that has not, in fact, been proven.

The following commands have been tested on an Ubuntu 22.04 system (which requires the `--disable-sandboxing`) and on a Debian Bookworm system. These commands download the HOL Light development package (rather than using the HOL Light package built into Bookworm), and should work on a wider range of Linux distributions, as long as the `apt` line is adapted appropriately.

```
sudo apt install opam wget -y

time opam init -a --disable-sandboxing
time opam switch create 4.05.0
eval `opam env`
time opam pin add camlp5 7.10 -y
time opam install num camlp-streams ocamlfind -y

git clone https://github.com/jrh13/hol-light
cd hol-light
git checkout 1a1de6ce7a6e9f60bec8bc501c426836d0e6b231
make

wget https://cr.yp.to/2023/lprrr-20230317.ml
time ocaml -I `camlp5 -where` camlp5o.cma -init hol.ml \
< lprrr-20230317.ml > lprrr-20230317.out
```

On the Ubuntu 22.04 system (with an AMD FX-8350 CPU), the timed commands were observed to take 39 seconds, 314 seconds, 72 seconds, 187 seconds, and 365 seconds respectively. The reader can check that the resulting `lprrr-20230317.out` file includes the definitions and theorems shown above, each certified by HOL Light to be a `thm`.