Markov chain quasi-Monte Carlo simulation using linear feedback shift register generators

Shin Harase

Ritsumeikan University

August 22, 2024

MCQMC 2024
(University of Waterloo)

This work was supported by JSPS KAKENHI Grant Numbers 22K11945, 18K18016.

Introduction: Markov chain QMC

Motivation:

- We consider the problem of estimating the expectation

$$E_\pi[f(\mathbf{X})] = \int_{\mathcal{X}} f(x)\pi(dx)$$

  using Markov chain Monte Carlo (MCMC) for a probability distribution π on a state space $\mathcal{X}$ and some function $f : \mathcal{X} \to \mathbb{R}$.
- We want to improve the accuracy by replacing IID uniform random points with quasi-Monte Carlo (QMC) points.
- Traditional QMC points (e.g., Sobol', Faure, Niederreiter–Xing, ...) are not straightforwardly applicable.
- Owen–Tribble (2005) and Chen–Dick–Owen (2011) proved that MCMC remains consistent if the driving sequences are completely uniformly distributed (CUD).
- However, the definition of CUD sequences is not constructive.

Introduction: Markov chain QMC

Motivation (continued):

- To obtain point sets that approximate CUD sequences, Chen-Matsumoto-Nishimura-Owen (2012) designed short-period Tausworthe generators (i.e., linear feedback shift register generators) optimized in terms of the equidistribution property, which is a coarse measure used in pseudorandom number generation.
- Harase (2021, 2024) designed short-period Tausworthe generators in terms of the t-value, which is a central measure in the theory of (t, m, s)-nets and (t, s)-sequences.

In this talk, we outline new Tausworthe generators for Markov chain QMC and present numerical experiments using Gibbs sampling.

References:

1. S. Harase, "A table of short-period Tausworthe generators for Markov chain quasi-Monte Carlo", J. Comput. Appl. Math. 384 (2021), 113136, 12 pp.
2. S. Harase, "A search for short-period Tausworthe generators over $\mathbb{F}_b$ with application to Markov chain quasi-Monte Carlo", J. Stat. Comput. Simul. 94 (2024), no. 9, 2040–2062.

CUD sequences

Definition (CUD sequences)

A one-dimensional sequence $u_0, u_1, u_2, u_3, u_4, \ldots \in [0,1)$ is said to be completely uniformly distributed (CUD) if the overlapping s-blocks

$$(u_i, u_{i+1}, \ldots, u_{i+s-1}) \in [0,1)^s, \quad i = 0, 1, 2, \ldots,$$

are uniformly distributed for every dimension $s \geq 1$.

- Then, the non-overlapping s-blocks $(u_{is}, u_{si+1}, \ldots, u_{(i+1)s-1})$ are also uniform (Chentsov, 1967), so we can use $\{u_i\}_{i=0}^{\infty}$ in this order.
- It is desirable that the s-blocks are highly uniform (Dick–Rudolf, 2014).

Chen et al. (2012) and Harase (2021, 2024) designed "approximate" CUD sequences

$$u_0, u_1, \ldots, u_{N-2}, u_{N-1} = u_0, \ldots \quad (N-1 : \text{period length}),$$

based on short-period Tausworthe generators that run for the entire-period.

Tausworhe generators (Tausworthe, 1965)

Let $\mathbb{F}_b$ be a finite field of prime power order b.
We define Tausworthe generators over $\mathbb{F}_b$ as polynomial LCGs:

$$\begin{aligned} \boldsymbol{X}_i(x) &= \boldsymbol{q}(x)\boldsymbol{X}_{i-1}(x) \mod \boldsymbol{p}(x), \\ \boldsymbol{X}_i(x)/\boldsymbol{p}(x) &= \boldsymbol{a}_{i\sigma}x^{-1} + \boldsymbol{a}_{i\sigma+1}x^{-2} + \boldsymbol{a}_{i\sigma+2}x^{-3} + \cdots \in \mathbb{F}_b((x^{-1})). \end{aligned}$$

Here, $\boldsymbol{p}(x), \boldsymbol{q}(x) \in \mathbb{F}_b[x]$ represent a modulus and multiplier. Assume

- $\boldsymbol{p}(x)$ is a primitive polynomial with degree $\boldsymbol{m}$;
- $\boldsymbol{q}(x)$ satisfies $\gcd(\sigma, b^m - 1) = 1$, where σ is a step size such that $\boldsymbol{q}(x) = x^{\sigma} \mod \boldsymbol{p}(x)$ and $0 < \sigma < b^m - 1$.

Let $\eta : \mathbb{F}_b \to \mathbb{Z}_b = \{0, 1, \ldots, b\}$ be a bijection with $\eta(0) = 0$. Then, we transform the formal power series into the w-digit output values

$$u_i = \sum_{j=0}^{w-1} \eta(a_{i\sigma+j}) \cdot b^{-j-1} \in [0, 1) \quad (w\text{: a digit number}).$$

In this setting, the output sequence $\{u_i\}_{i=0}^{\infty}$ attains the maximal period $b^m - 1$. We assume the maximal periodicity and $w \geq m$.

Tausworhe generators (Tausworthe, 1965)

Example: $b = 3$ (base). $w = 3$ (digit number). $\eta = \mathrm{id}$.
$p(x) = x^3 + 2x^2 + x + 1 \Leftrightarrow a_i = -2a_{i-1} - a_{i-2} - a_{i-3}$.
$q(x) = x^2 + 2x (= x^5 \mod p(x))$, so the step size $\sigma = 5$.

$$00110102122210022020121112001101021222100 2 \cdots \in \mathbb{F}_3$$

$$\begin{aligned}
X_0(x)/p(x) &= 0x^{-1} + 0x^{-2} + 1x^{-3} + \cdots \mapsto u_0 = 0.001_{(3)} \\
X_1(x)/p(x) &= 1x^{-1} + 0x^{-2} + 2x^{-3} + \cdots \mapsto u_1 = 0.102_{(3)} \\
X_2(x)/p(x) &= 2x^{-1} + 2x^{-2} + 1x^{-3} + \cdots \mapsto u_2 = 0.220_{(3)} \\
&\vdots
\end{aligned}$$

Let $N := b^m$. We run Tausworthe generators for the entire period and construct the s-dimensional overlapping points
$\mathbf{u}_0 = (u_0, \ldots, u_{s-1}), \mathbf{u}_1 = (u_1, \ldots, u_s), \ldots, \mathbf{u}_{N-2} = (u_{N-2}, u_0, \ldots, u_{s-2})$. Adding the origin $\{\mathbf{0}\}$, we regard a point set

$$P_s = \{\mathbf{0}\} \cup \{\mathbf{u}_i\}_{i=0}^{N-2} \qquad (|P_s| = b^m)$$

as a QMC point set, which corresponds to a polynomial Korobov lattice.

(t, m, s)-nets and t-values

In the study of QMC, the t-value is widely used as a measure of uniformity.

Definition ((t, m, s)-nets and t-values)

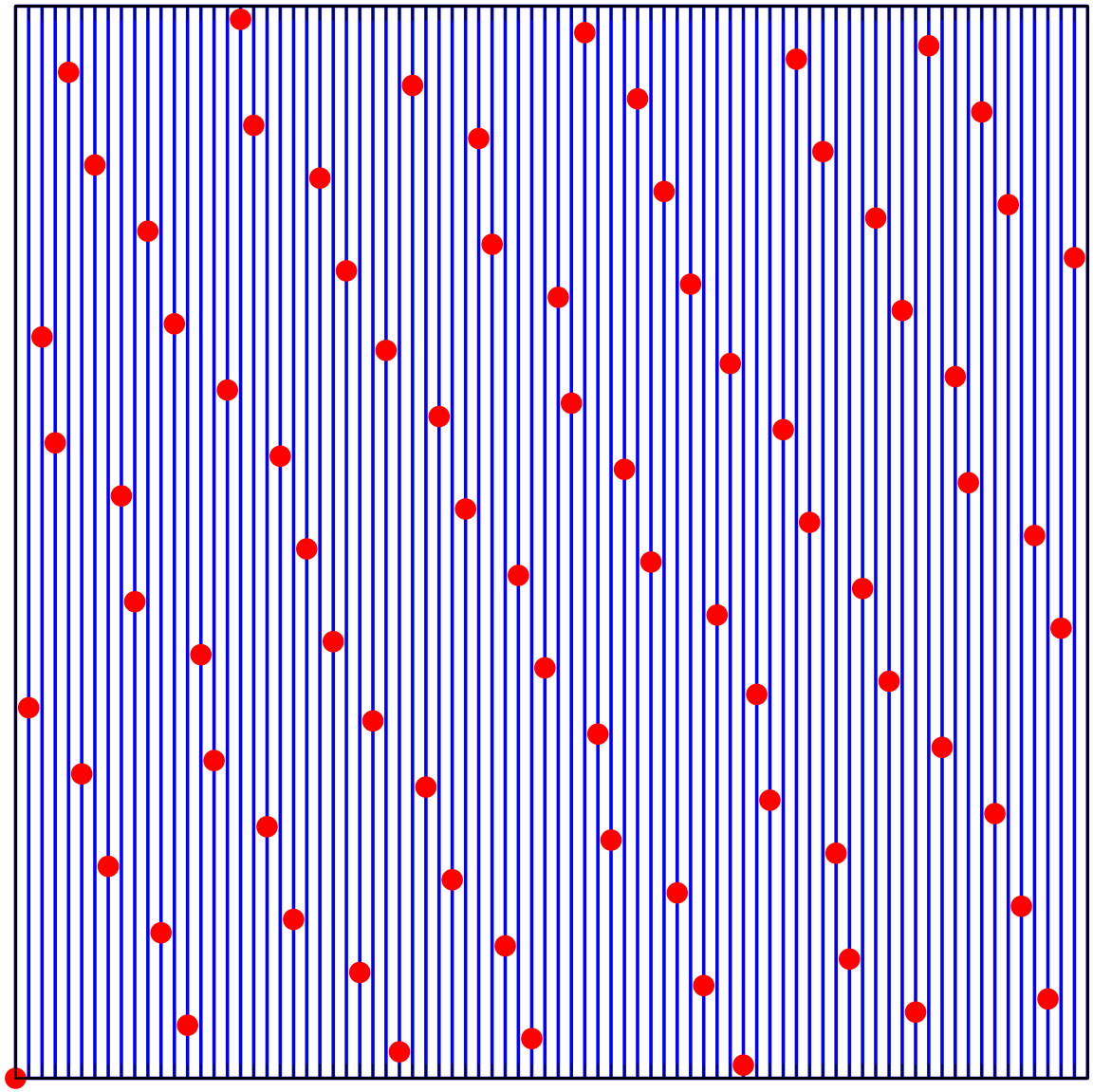
- Let $s \geq 1$ be a dimension. Let t be an integer with $0 \leq t \leq m$.
- Let E be a b-adic box $E = \prod_{j=1}^{s} \left[\frac{l_j}{b^{d_j}}, \frac{l_j+1}{b^{d_j}} \right) \subset [0,1)^s$.
- Let $P_s \subset [0,1)^s$ be a point set consisting of $N = b^m$ points.

P_s is a (t, m, s)-net in base b $\overset{\text{def}}{\iff}$ Every b-adic box E of volume b^{t-m} contains exactly b^t points of P_s.
The smallest t for which P_s is a (t, m, s)-net is called the t-value.

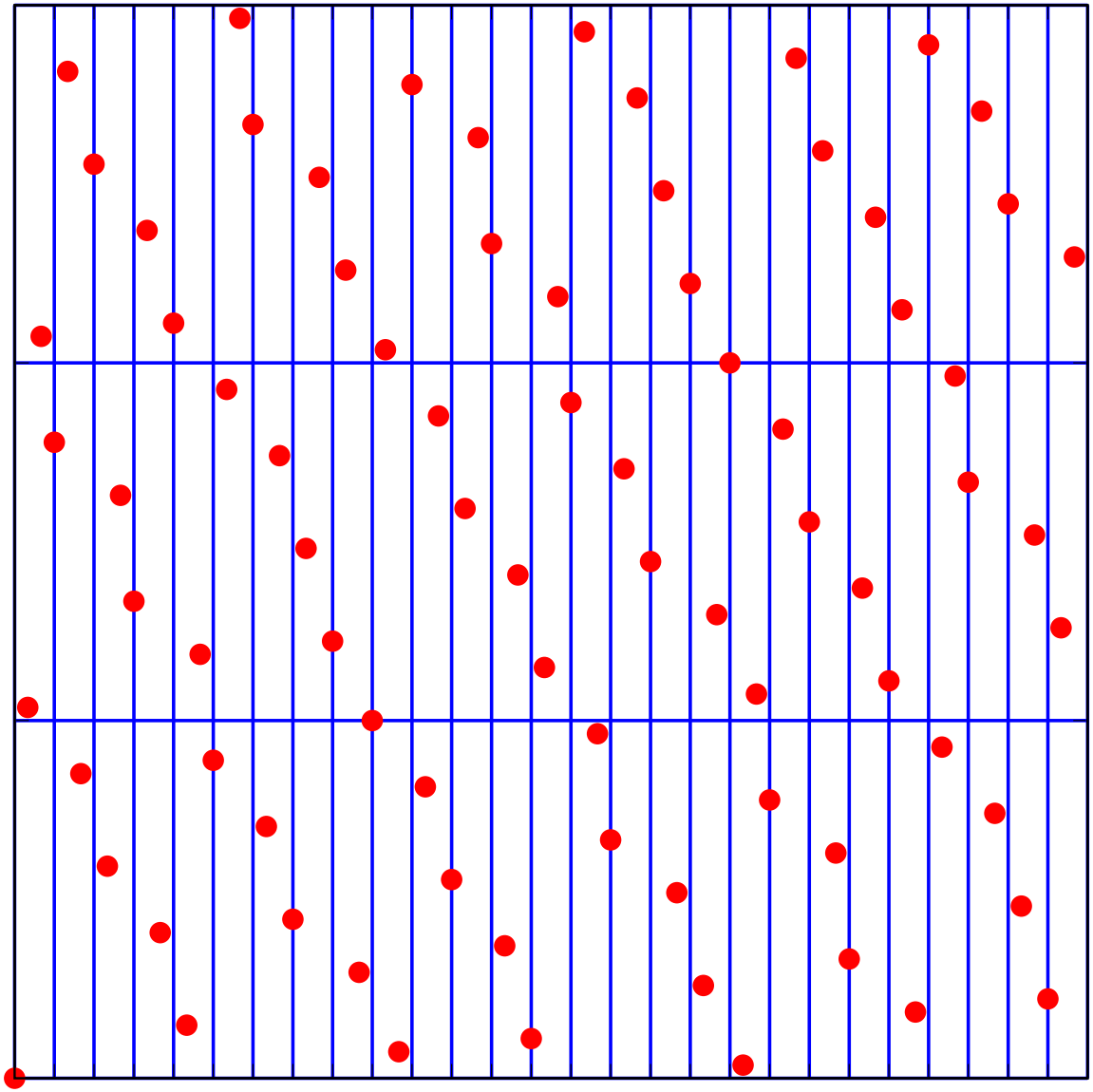
To construct a point set that approximates the CUD property, Harase (2021, 2024) searched a pair of polynomials $(p(x), q(x))$ whose t-values are small for each $s = 1, 2, 3, \ldots$, that is, Tausworthe generators

1. over $\mathbb{F}_2$ with t-values zero for $s = 2$ and small for $s \geq 3$; and
2. over $\mathbb{F}_4$ with t-values zero up to $s = 3$ and small for $s \geq 4$.

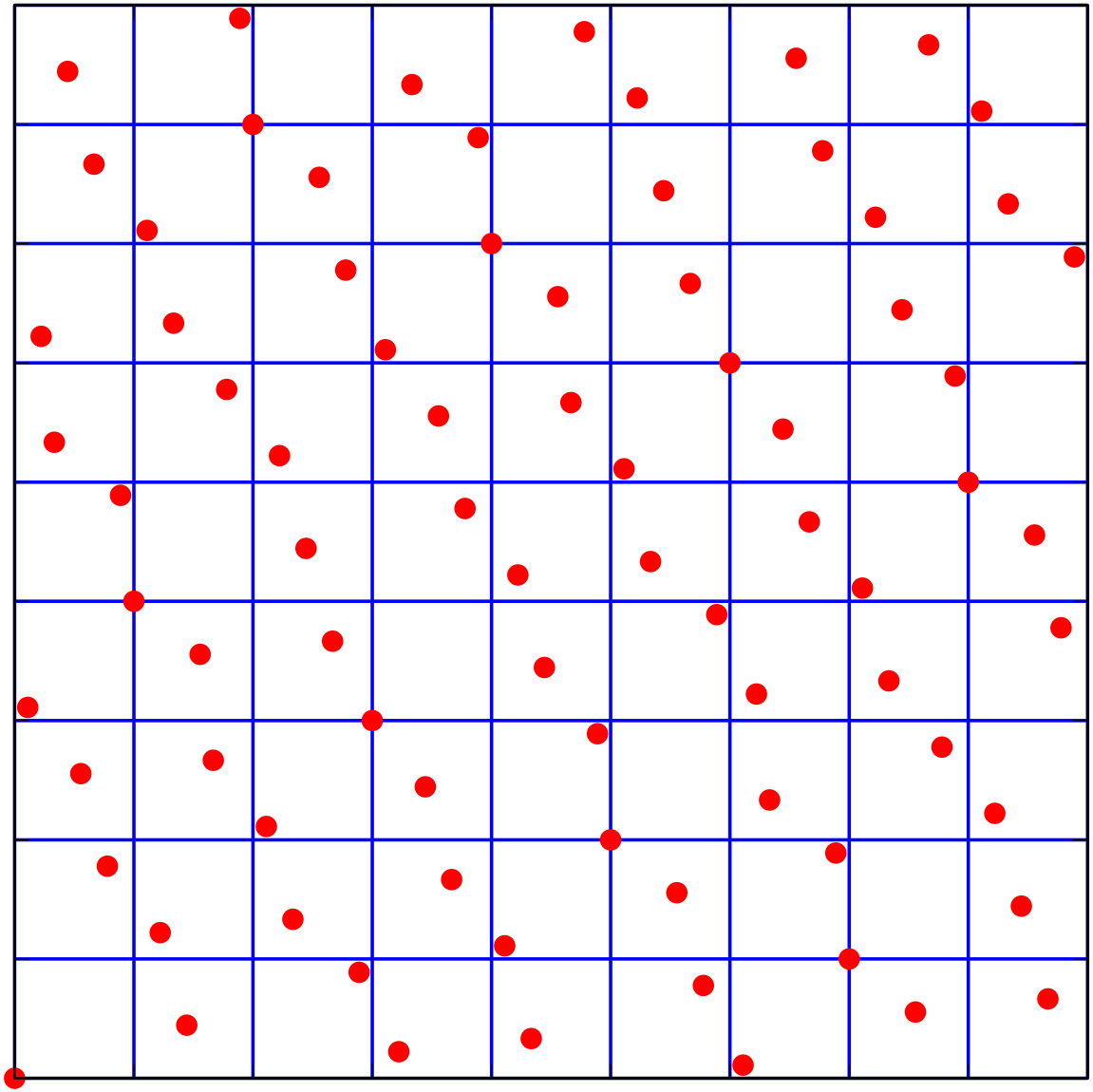
Example: $(0, 4, 2)$-net in base 3



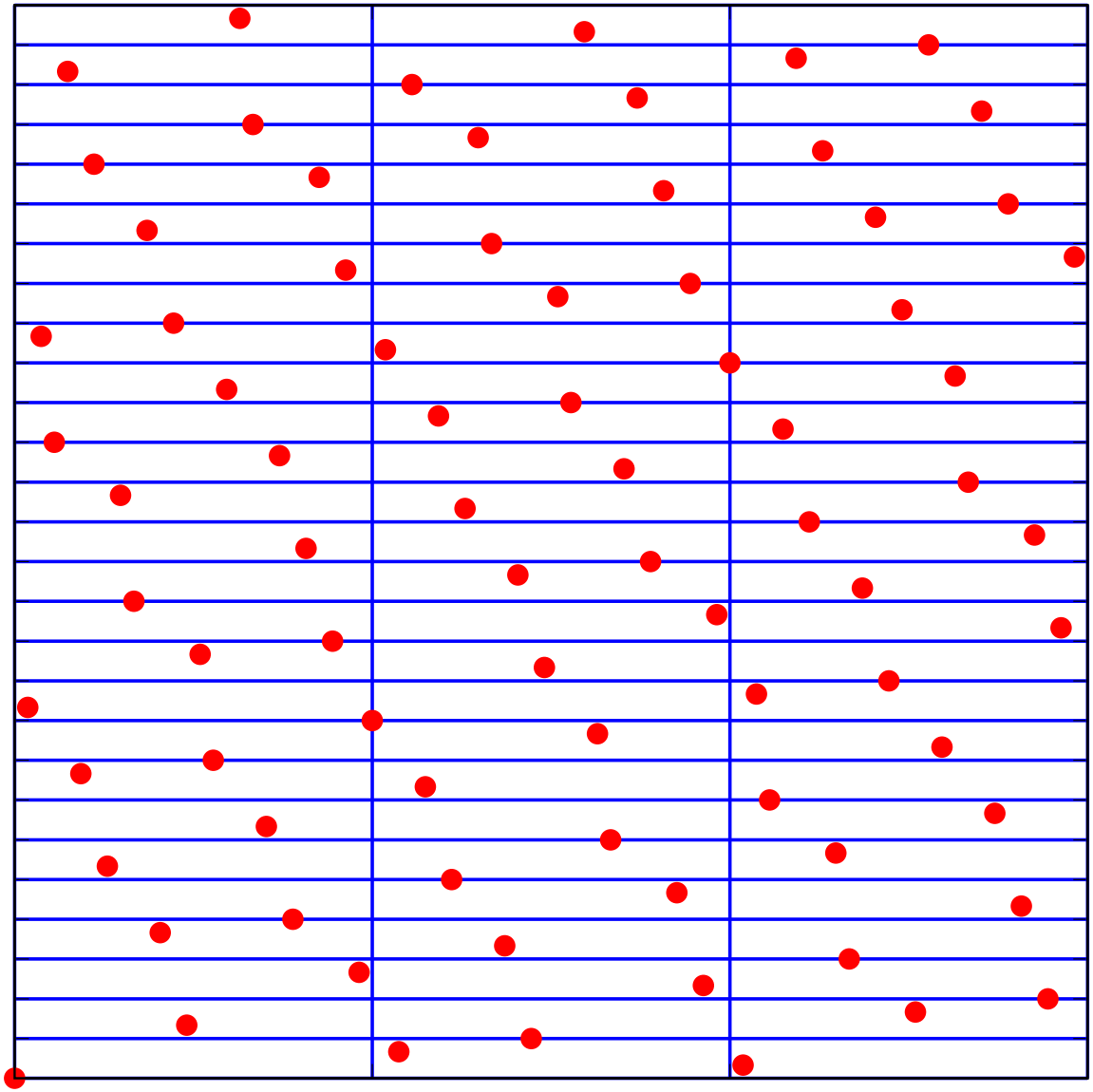
Example: $(0, 4, 2)$-net in base 3



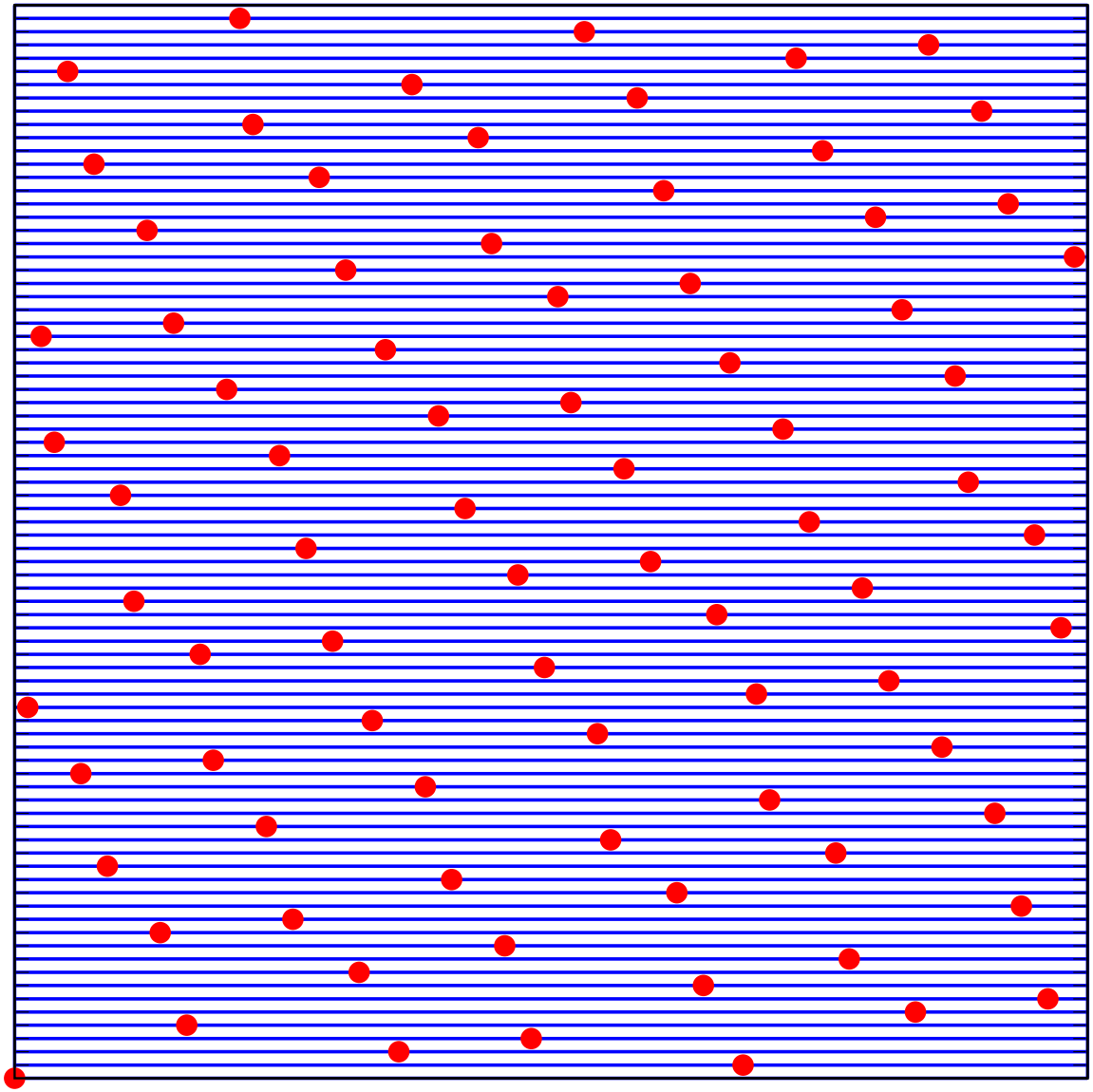
Example: $(0, 4, 2)$-net in base 3



Example: $(0, 4, 2)$-net in base 3



Example: $(0, 4, 2)$-net in base 3



Search algorithm

We briefly introduce the search algorithm (Harase, 2021 and 2024).

Theorem (Niederreiter, 1992) and (Tezuka–Fushimi, 1993)

Let $p(x)$ and $q(x) \in \mathbb{F}_b[x]$ be a modulus and multiplier of Tausworthe generators. Then, the 2-dimensional point set P_2 attains the t-value zero (in base b) if and only if the partial quotients in the continued fraction of $q(x)/p(x)$ are all of degree 1.

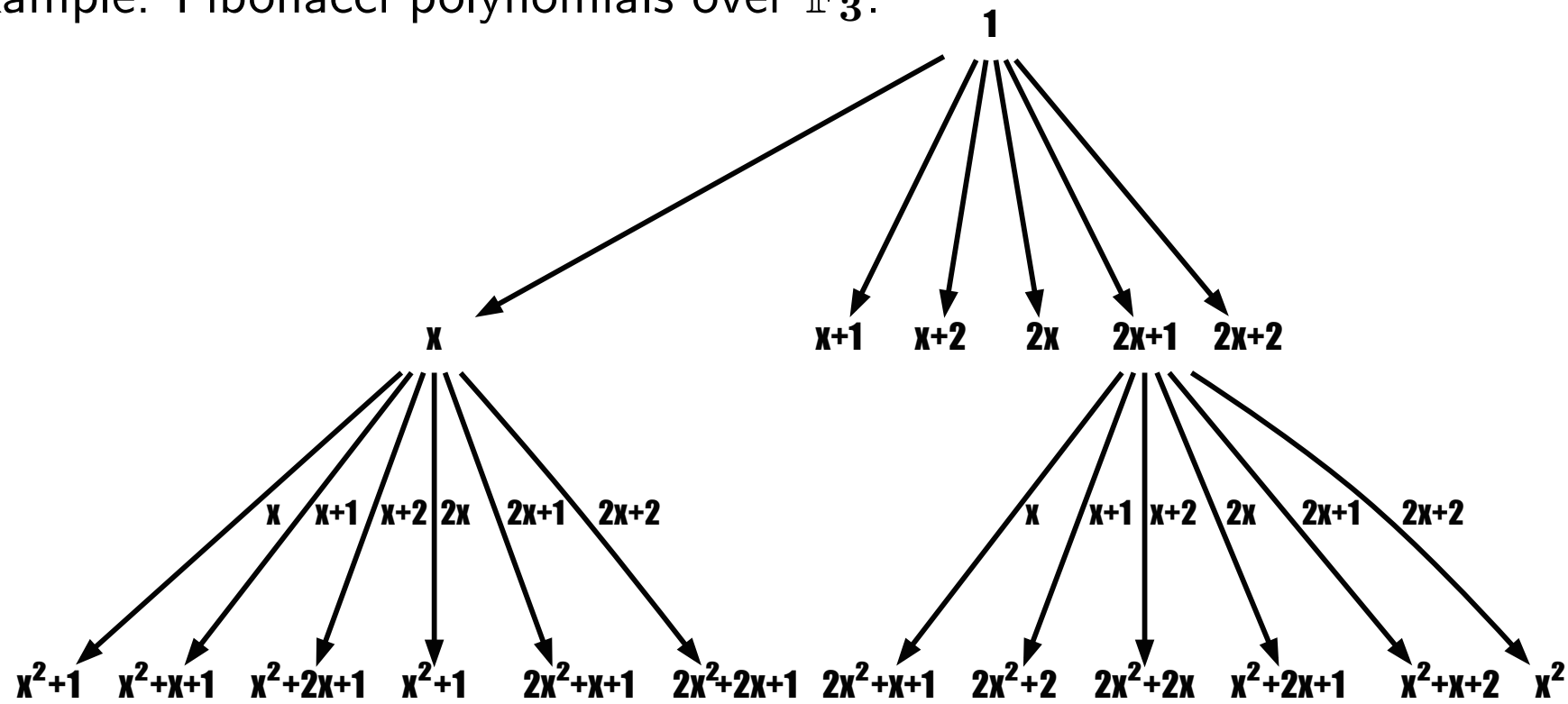
To obtain such pairs $(p(x), q(x))$, we consider a polynomial analogue of Fibonacci numbers over $\mathbb{F}_b$ (cf., Tezuka and Fushimi (1993)):

$$
\begin{aligned}
&F_k(x) = A_k(x)F_{k-1}(x) + F_{k-2}(x), \quad k = 1, 2, \ldots, m, \\
&F_{-1}(x) = 0, F_0(x) = 1, \\
&A_k(x) = \beta x + \gamma, \quad \beta \in \mathbb{F}_b^* \text{ and } \gamma \in \mathbb{F}_b.
\end{aligned}
$$

Then, the partial quotients in the continued fraction of $F_{m-1}(x)/F_m(x)$ are all of degree 1, i.e., $F_{m-1}(x)/F_m(x) = [0; A_m, A_{m-1}, \cdots, A_1]$.

Fibonacci polynomials over $\mathbb{F}_b$ and tree structures

Example: Fibonacci polynomials over $\mathbb{F}_3$.



$$F_k(x) = A_k(x)F_{k-1}(x) + F_{k-2}(x), \quad k = 1, 2, \ldots, m$$
$$F_{-1}(x) = 0, F_0(x) = 1$$
$$A_k(x) = \beta x + \gamma, \quad \beta \in \mathbb{F}_b^* \text{ and } \gamma \in \mathbb{F}_b.$$

From all the pairs $(F_m(x), F_{m-1}(x))$, we choose a suitable pair $(p(x), q(x))$ with t-values zero for $s = 2$ and small for $s \geq 3$.

Comparison of the t-values: our new vs existing results

A search for Tausworthe generators over $\mathbb{F}_2$ (Harase, 2021):

$m = 16$

dim. s	2	3	4	5	6	7	8	9	10	11	12	13	14	15
New	0	3	4	7	7	8	10	10	10	11	11	11	11	11
Existing	3	4	5	8	8	8	8	8	10	10	10	10	10	10

(Existing: Tausworthe generators developed by Chen et al. (2012).)

A search for Tausworthe generators over $\mathbb{F}_4$ (Harase, 2024):

$m \backslash s$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
5	0	0	0	1	1	2	2	2	2	2	2	2	2	2	2
6	0	0	0	1	2	2	2	2	3	3	3	3	3	3	3
7	0	0	0	1	2	2	2	3	3	3	3	3	3	4	4
8	0	0	0	1	2	4	4	4	4	4	4	4	4	4	4
9	0	0	0	1	3	3	3	3	3	4	4	4	4	4	4
10	0	0	0	2	2	3	3	3	4	4	4	5	5	6	6

Numerical example: 3-dimensional Gaussian Gibbs

We provide numerical examples to confirm the performance of Markov chain QMC.

We consider the three-dimensional Gaussian (normal) distribution $\mathcal{N}(\mu, \Sigma)$, where

$$\mu = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0.3 & -0.2 \\ 0.3 & 1 & 0.5 \\ -0.2 & 0.5 & 1 \end{pmatrix}.$$

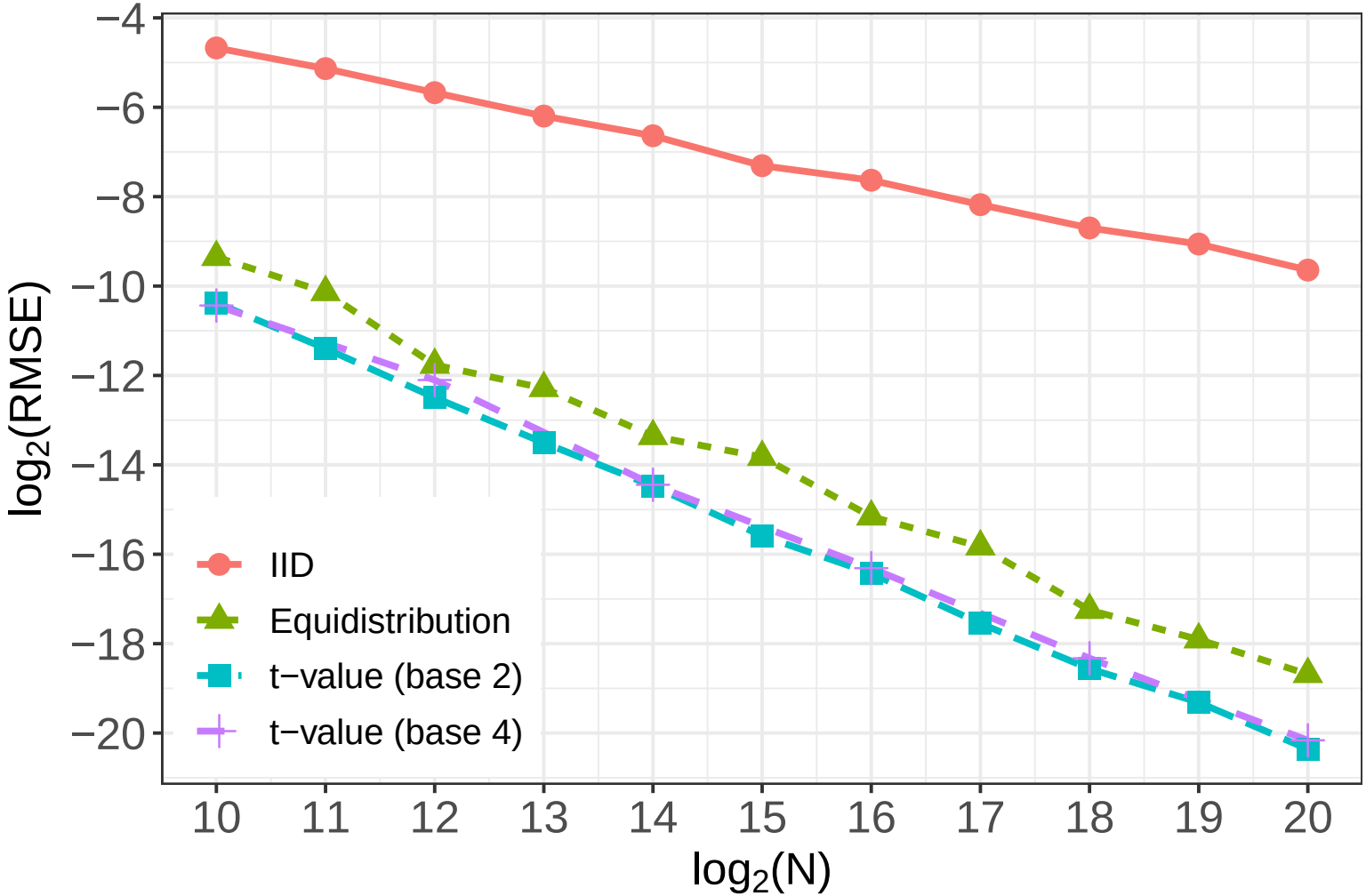
Then, a Gibbs sampling scheme can be implemented as the iteration of the calculation of one-dimensional normal distribution as follows:

$$X_k \mid \mathbf{X}_{-k} \sim \mathcal{N}(\mu_k + \Sigma_{k,-k}\Sigma_{-k,-k}^{-1}(\mathbf{X}_{-k} - \mu_{-k}), \Sigma_{k,k} - \Sigma_{k,-k}\Sigma_{-k,-k}^{-1}\Sigma_{-k,k})$$

for $k = 1, 2, 3$.

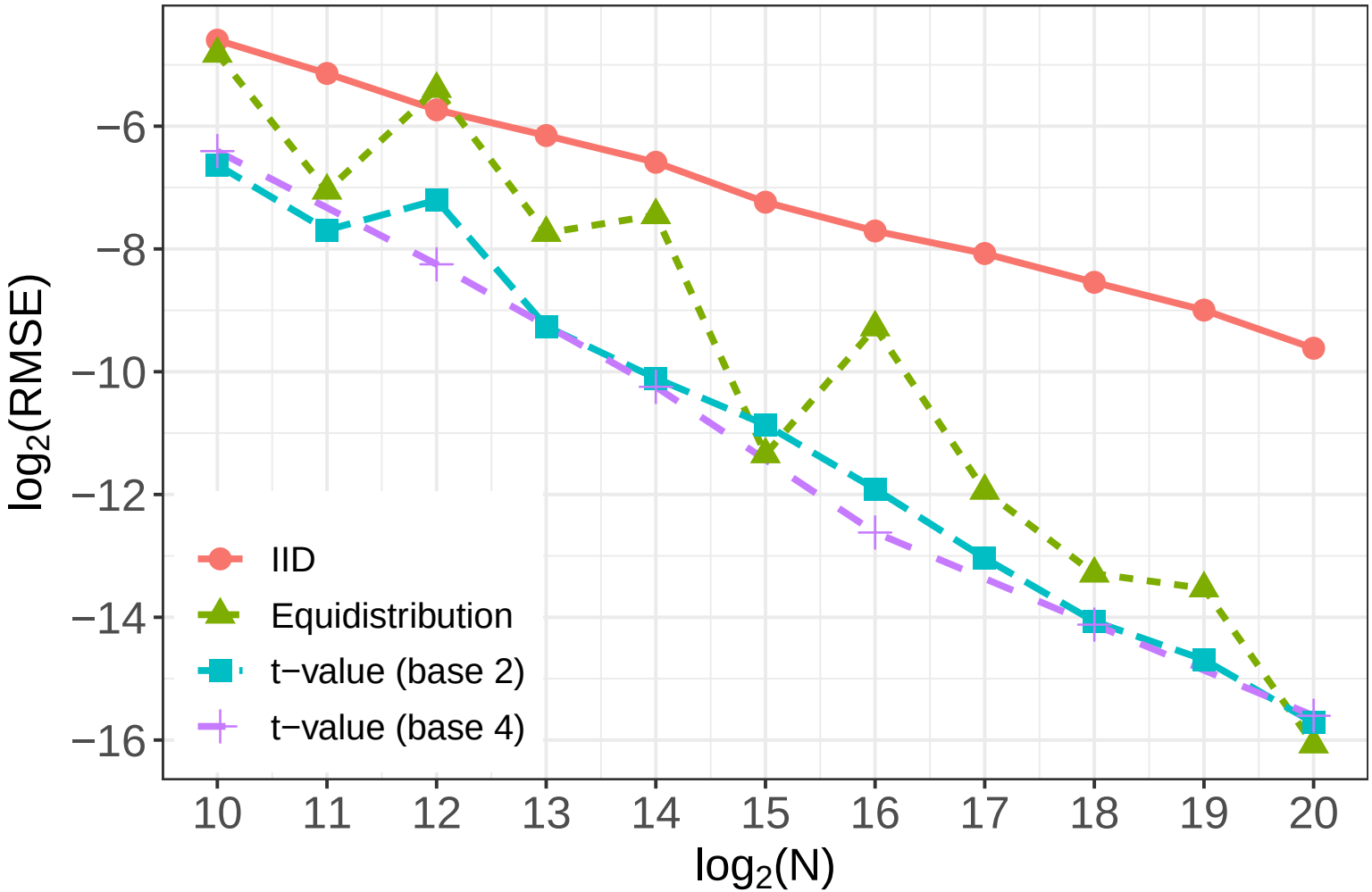
Numerical example: 3-dimensional Gaussian Gibbs

We estimate $E[X_1]$ with true value 0 by taking the sample mean.
We calculate RMSEs using 300 digital shifts.



Numerical example: 3-dimensional Gaussian Gibbs

We estimate $E[X_1X_2X_3]$ with true value 0 by taking the sample mean.
We calculate RMSEs using 300 digital shifts.



Numerical example: Bayesian linear regression

We use the Boston housing data.
Harrison and Rubinfeld (1978) built a linear regression model given by

$$\begin{aligned}\log(\mathrm{MEDV}) = {} & \beta_0 + \beta_1 \mathrm{CRIM} + \beta_2 \mathrm{ZN} + \beta_3 \mathrm{INDUS} + \beta_4 \mathrm{CHAS} + \beta_5 \mathrm{NOX}^2 \\ & + \beta_6 \mathrm{RM}^2 + \beta_7 \mathrm{AGE} + \beta_8 \log(\mathrm{DIS}) + \beta_9 \log(\mathrm{RAD}) + \beta_{10} \mathrm{TAX} \\ & + \beta_{11} \mathrm{PTRATIO} + \beta_{12} \mathrm{B} + \beta_{13} \log(\mathrm{LSTAT}) + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \tau^2),\end{aligned}$$

where

- The housing price MEDV is a response variable;
- The constant term, CRIM, . . . , LSTAT are 14 explanatory variables.

We now consider Bayesian inference (e.g., see Hoff (2009)).

- Assume that $\beta = (\beta_0 \ldots, \beta_{13})$ and τ^2 are independent and follow the normal and inverse-gamma prior distributions.
- We calculate the posterior mean estimates $E[\beta]$ and $E[\tau^2]$ by running the Gibbs sampler for $N = 2^{16}$ iterations after a burn-in period.

Numerical example: Bayesian linear regression

The following table shows a summary of sample variances of posterior mean estimates $E[\beta_i]$ and $E[\tau^2]$ using 300 digital shifts.

$N = 2^{16}$					
Parameter	β_0	β_1	β_2	β_3	β_4
IID	3.07e-07	2.74e-11	4.13e-12	8.43e-11	1.64e-08
Equidistribution	9.28e-12	6.59e-16	1.07e-16	2.37e-15	4.71e-13
t-value (base 2)	1.40e-12	9.12e-17	1.82e-17	3.55e-16	1.71e-13
t-value (base 4)	1.73e-12	9.86e-17	1.71e-17	4.20e-16	7.31e-14
Parameter	β_5	β_6	β_7	β_8	β_9
IID	1.72e-07	2.22e-11	4.26e-12	1.54e-08	5.87e-09
Equidistribution	4.73e-12	6.69e-16	1.14e-16	4.70e-13	1.67e-13
t-value (base 2)	7.11e-13	7.65e-17	8.86e-18	4.47e-14	1.42e-14
t-value (base 4)	9.01e-13	1.24e-16	1.63e-17	7.71e-14	3.20e-14
Parameter	β_{10}	β_{11}	β_{12}	β_{13}	τ^2
IID	2.25e-13	3.48e-10	1.89e-13	1.09e-08	7.35e-11
Equidistribution	6.91e-18	1.08e-14	4.52e-18	2.34e-13	2.36e-15
t-value (base 2)	1.18e-18	1.36e-15	6.42e-19	3.02e-14	9.24e-16
t-value (base 4)	1.36e-18	2.28e-15	1.20e-18	4.68e-14	1.00e-15

Conclusion

Summary:

- For Markov chain QMC, we conducted a search of short-period Tausworthe generators in terms of the t-value, that is, Tausworthe generators
  1. over $\mathbb{F}_2$ with t-values zero for $s = 2$ and small for $s \geq 3$; and
  2. over $\mathbb{F}_4$ with t-values zero up to $s = 3$ and small for $s \geq 4$.
- In the parameter search, we used Fibonacci polynomials over $\mathbb{F}_b$.
- We reported numerical examples using Gibbs sampling in which our new Tausworthe generators perform comparable to or even better than the existing Tausworthe generators (Chen et al., 2012).

References:

1. S. Harase, "A table of short-period Tausworthe generators for Markov chain quasi-Monte Carlo", J. Comput. Appl. Math. 384 (2021), 113136, 12 pp.
2. S. Harase, "A search for short-period Tausworthe generators over $\mathbb{F}_b$ with application to Markov chain quasi-Monte Carlo", J. Stat. Comput. Simul. 94 (2024), no. 9, 2040–2062.

The code in C is available at `https://github.com/sharase`.