

Received 16 November 2023, accepted 1 December 2023, date of publication 12 December 2023,
date of current version 21 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3341422

RESEARCH ARTICLE

Correlation-Diversified Portfolio Construction by Finding Maximum Independent Set in Large-Scale Market Graph

RYO HIDAKA^{id}, YOHEI HAMAKAWA^{id}, JUN NAKAYAMA^{id}, AND
KOSUKE TATSUMURA^{id}, (Member, IEEE)

Corporate Research and Development Center, Toshiba Corporation, Kawasaki 212-8582, Japan

Corresponding author: Ryo Hidaka (ryo.hidaka@toshiba.co.jp)

ABSTRACT Correlation-diversified portfolios can be constructed by finding the maximum independent sets (MISs) in market graphs with edges corresponding to correlations between two stocks. The computational complexity of finding the MIS increases exponentially as the size of the market graph increases, making the MIS selection in a large-scale market graph difficult. Here we construct a diversified portfolio through solving the MIS problem for a large-scale market graph with a combinatorial optimization solver (an Ising machine) based on a quantum-inspired algorithm called simulated bifurcation (SB) and investigate the investment performance of the constructed portfolio using long-term historical market data at the Tokyo Stock Exchange. Comparisons using stock universes of various sizes (TOPIX 100, Nikkei 225, TOPIX 1000, and TOPIX which includes approximately 2,000 constituents) show that the SB-based solver outperforms conventional MIS solvers in terms of computation-time and solution-accuracy. By using the SB-based solver, we optimized the parameters of a MIS portfolio strategy through iteration of the backcast simulation that calculates the performance of the MIS portfolio strategy based on a large-scale universe covering more than 1,700 Japanese stocks for a long period of 10 years. It has been found that the best MIS portfolio strategy (Sharpe ratio = 1.16, annualized return/risk = 16.3%/14.0%) outperforms the major indices such as TOPIX (0.66, 10.0%/15.2%) and MSCI Japan Minimum Volatility Index (0.64, 7.7%/12.1%) for the period from 2013 to 2023. Factor analyses reveal that the selection of small-capitalization and low-correlation stocks results in the portfolio performance with not only the relatively low risk (the diversification effect as expected) but also the relatively high return.

INDEX TERMS Portfolio construction, market graph, maximum independent set, combinatorial optimization, FPGA, Ising machine, simulated bifurcation, quantum-inspired.

I. INTRODUCTION

From the establishment of Modern Portfolio Theory (MPT) by Harry Markowitz and William F. Sharpe in the mid-20th century to the present [1], [2], improving the risk-return characteristics of portfolios by diversifying across multiple financial instruments has remained an important issue for both institutional and individual investors. If portfolios can be constructed with stocks that have low correlations to

each other, the price movements of the component stocks will be disparate and their contributions to the overall portfolio performance will cancel each other out, thus reducing the risk of the portfolio. Such a correlation effect is one of the basis of the MPT. In the last decade or two, attention has focused on risk-based portfolios and minimum volatility indices [3], and a variety of indices and funds linked to them have been launched. There are also a lot of studies on the portfolios that focus on risk, such as ones using genetic algorithm [4] and deep reinforcement learning [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Harikrishnan Ramiah^{id}.

The “market graph,” which represents the correlation of stocks, has been actively studied [6], [7], [8], [9], [10]. The market graph is an undirected graph with stocks as nodes, connected by edges between nodes (stocks) with a certain level of correlation; in other words, its graph structure (connection relationship) represents the correlation of the entire stock universe. A portfolio constructed with unconnected nodes on the market graph, or an independent set, will be a diversified portfolio with low correlations to each other, which is expected to reduce risk. Kalyagin et al. [9] investigated the relationship of independent sets of market graphs to Markowitz’s portfolio theory and demonstrated the usefulness of the market graph.

The independent set of a market graph with the maximum number of nodes is obtained by solving the maximum independent set (MIS) problem, one of the combinatorial optimization problems. While the MIS problem has a variety of practical applications [11], [12], [13], [14], it is known to be nondeterministic polynomial time (NP)-hard, hence the computational complexity increases exponentially as the number of nodes (stocks in the market graph) increases. Butenko [6] used a greedy heuristic to solve the MIS problem for market graphs in a reasonable time.

Recently, Ising machines [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], which specialize in obtaining high-accuracy solutions of combinatorial optimization problems in short times, have attracted much attention. The Ising machine is a device that searches for the ground-state of the Ising spin model (this search is called the Ising problem), i.e., a mathematical model of magnetic materials that consists of binary variables, called spins, coupled with each other by pairwise interactions. Ising machines with full spin-to-spin connectivity are applicable, without embedding the target problems into locally-connected native graphs, to various graph analyses such as partitioning (inc. maximum clique or maximum independent set problems), clustering, coloring, and path search [34] and also to various graph structures including special structures such as modular graphs [35]. The generalization of the Ising machine (quadratic) to higher-order problems [36], [37] has also been studied to analyze higher-order networks [38], [39], [40], [41], [42], [43].

Simulated bifurcation (SB) [15] is a quantum-inspired heuristic algorithm for combinatorial optimization. The algorithm of SB is highly parallelizable and enables developing the massive-parallel accelerators (Ising machines) at a large scale (including more than several thousand Ising spins) with parallel processors such as field programmable gate arrays (FPGAs) [15], [16], [17], [18] and graphics processing units (GPUs) [15], [17]. SB was derived from a classical counterpart to a quantum adiabatic optimization method called quantum bifurcation [45], and variations of SB have been proposed: adiabatic SB (aSB) [15], ballistic SB (bSB) [17], discrete SB (dSB) [17], and heated SB [19]. The hardware implementations of SBs are called

simulated bifurcation machines (SBMs). Financial automated trading systems using SBMs for detecting short-lived trading opportunities have been reported [46], [47], [48].

Yarkoni et al. [44] solved a random MIS problem with D-Wave’s quantum annealing (QA) processor [20], [21], an Ising machine based on quantum principles, and reported speed and solution-accuracy advantages over conventional MIS solvers at graph sizes of up to 40 nodes. The number of nodes (stocks) of 40 is, however, not sufficient for practical portfolio selection, as it requires handling thousands of stocks when considering the whole constituents of a market index as the universe for stock selection [e.g. the Tokyo Stock Price Index (TOPIX) as the proxy for the Japanese stock market involves more than 2,000 stocks, and the National Association of Securities Dealers Automated Quotations (NASDAQ) Composite Index for the NASDAQ market involves more than 3,000 stocks]. The characteristic of MIS portfolio strategy for a large-scale universe and for a long period has not been sufficiently investigated.

In this work, we build an FPGA-based accelerator for bSB (hereafter, the SBM), and by using it, investigate the performance of a MIS portfolio strategy based on the market graphs covering over 1,700 Japanese stocks when varying the strategy parameters by iterating the 10-year backcast simulation. The performance of the best MIS portfolio strategy is compared with the Japanese major indices from the perspective of risk-return characteristics.

The SBM supports fully-connected 2,048-spin Ising models and is capable of solving 2,048-node MIS problems, which is implemented as a peripheral component interconnect-express (PCIe)-attachable, look-aside type, FPGA-based acceleration card to the CPU-based host system. The backcast simulation is carried out with the CPU, but the computationally-hard parts, namely the parts to solve the MIS problems, are offloaded to the SBM without large network-communication overheads. We also developed the hardware abstraction layer (HAL) for the SBM and APIs (Application Programming Interface) for C/C++ and Python programming languages, making the SBM accessible to the backcast simulator.

To evaluate SBM performance as a MIS solver, we compare it with conventional MIS solvers, NetworkX [49] and OR-tools [50], in terms of computation-time and solution-accuracy when solving the MIS problems of market graphs representing the constituents of Japanese stock indices (TOPIX, TOPIX 1000, Nikkei 225, and TOPIX 100). See APPENDIX A for the details of the major market indices in the Tokyo Stock Exchange (TSE).

We repeat the backcast simulations of the MIS portfolio strategy with TOPIX as the stock universe (over 1,700 constituent stocks) for the period from 2013 to 2023 under a total of 38 parameter settings (a combination of 19 patterns of thresholds that determine the edge connections of the market graphs and two different asset-allocation methods for the selected portfolios). And then, we compare the performance

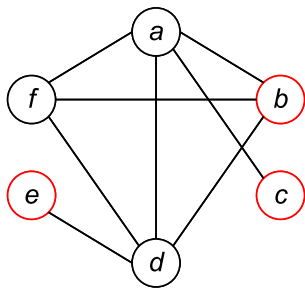


FIGURE 1. Market graph for an N -stock universe ($N = 6$). Red nodes are the maximum independent set (MIS).

of the best MIS portfolio strategy with the major indices such as TOPIX and Morgan Stanley Capital International (MSCI) Japan Minimum Volatility Index. We show a superior risk-return characteristic for the MIS strategy and furthermore analyze the factors contributing to the risk-return characteristic.

II. COMPOSITION OF SECTIONS

This paper is across several fields including financial engineering, discrete mathematics, and computer engineering. The rest of the paper is organized as follows. Section III (Portfolio strategy) describes the MIS portfolio strategy and its formulation. Section IV (System architecture) details the MIS portfolio simulator and the interface and core circuit of the SBM. Section V (MIS solver performance) compares the SB-based MIS solver with the conventional MIS solvers. Section VI (Portfolio performance) systematically shows the performance of the portfolio strategy when varying the strategy parameters and analyzes the property of the best MIS strategy. Section VII concludes the paper.

III. PORTFOLIO STRATEGY

A. PORTFOLIO CONSTRUCTION AND MANAGEMENT

1) PORTFOLIO SELECTION

The proposed strategy selects a portfolio of stocks from a stock universe (a group of tradable stocks) according to the maximum independent set (MIS) found in the market graph representing the stock universe. The market graph is an undirected graph with nodes corresponding to the stocks in the universe and each edge corresponding to the presence of the correlation between the two nodes (/stocks) connected by the edge. The definition of the market graph and the constituents of the stock universes examined in this work will be detailed in Sec. III-B1 and Sec. III-B3, respectively. Fig. 1 shows an example of the market graph for an N -stock universe ($N = 6$) and the MIS in there. The solution to the MIS problem is a binary vector, where i th binary value means whether the i th stock is involved in the MIS. In this strategy, we choose the stocks involved in the MIS as the constituents of the portfolio.

2) PORTFOLIO COMPOSITION

After the portfolio selection, we determine the portfolio composition, i.e., the asset allocation to each constituent of

the portfolio. The weight w_i represents the ratio of the asset allocated for the stock i to the total asset. We examine the following two asset-allocation methods.

A) Equal Weight (EW): All constituents have the same weights; the weight w_i of stock i is represented by

$$w_i = \frac{1}{N}. \quad (1)$$

B) Inverse Volatility Weight (IVW): The weight is proportional to the inverse of the volatility of each constituent; w_i is defined by

$$w_i = \frac{v_i^{-1}}{\sum_{k=1}^N v_k^{-1}}, \quad (2)$$

where the volatility v_i is the standard deviation of the logarithmic daily return $[R_i(t)]$ [51]. The $R_i(t)$ is expressed by

$$R_i(t) = \ln \frac{P_i(t)}{P_i(t-1)}, \quad (3)$$

where $P_i(t)$ is the closing price of stock i on the day t (the last traded price on the business day) and the dividend (if exists) is considered. The standard deviation of $R_i(t)$ over the last three years is the volatility v_i .

3) PORTFOLIO REBALANCE

The portfolio is rebalanced monthly; the portfolio selection and weighting are carried out every month (at the end of each month). For the difference between the new portfolio and the last-month portfolio in terms of the constituents and composition, we take into account the trading costs corresponding to 0.1% of the total amount of transactions (the sum of the amount of shares purchased and the absolute amount of shares sold). The evaluation value of a portfolio in month t , $P_{PF}(t)$, is the sum of the evaluation values of constituents in the portfolio, where the evaluation value of a constituent is the product of the base price and share at the end of the month. The monthly return of the portfolio strategy, $R_{PF}(t)$, is expressed by

$$R_{PF}(t) = \frac{P_{PF}(t)}{P_{PF}(t-1)} - 1. \quad (4)$$

B. FINDING MIS IN MARKET GRAPH

1) DEFINITION OF MARKET GRAPH

We follow the definition and generation procedure of the market graph in [6]. In the correlation matrix C , the correlation coefficient $C_{i,j}$ between stocks i and j is calculated for the last T days as follows.

$$C_{i,j} = \frac{\sum_t (R_i(t) - \bar{R}_i)(R_j(t) - \bar{R}_j)}{\sqrt{\sum_t (R_i(t) - \bar{R}_i)^2} \sqrt{\sum_t (R_j(t) - \bar{R}_j)^2}}. \quad (5)$$

$R_i(t)$ is the logarithmic daily return defined by Eq. 3. The average of $R_i(t)$ for the last T days, \bar{R}_i , is given by

$$\bar{R}_i = \frac{1}{T} \sum_t^T R_i(t). \quad (6)$$

Upon the portfolio selection at the beginning of each month, we recalculate the correlation matrix C using historical price data for the last three years (the T days are the business days involved in the last three years). The correlation coefficient $C_{i,j}$ varies from -1 to 1. We can also consider $C_{i,j}$ as the similarity of the price fluctuations of the stocks i and j , i.e., the greater the value, the more similar the price movements of the two stocks.

The connectivity (or topology) of the market graph is determined by the correlation matrix C and a threshold parameter of θ ; the nodes i and j in the market graph are connected if the correlation coefficient $C_{i,j}$ is equal to or greater than θ . The connectivity of the market graph is sensitive to θ and the choice of θ characterizes the performance of a MIS portfolio strategy. In this work, we determine the best θ by analyzing the long-term performance of the MIS portfolio strategy for one of various θ and then assume using a fixed θ for the MIS strategy (not changed monthly).

2) QUBO FORMULATION OF MIS PROBLEM

We describe the MIS problem in the form of quadratic unconstrained binary optimization (QUBO) according to [34]. The QUBO problem is mathematically equivalent to the Ising problem. Considering a graph $G = (V, E)$ with a node set V and an edge set E , the connection $f_{i,j}$ between nodes i and j is represented as follows.

$$f_{i,j} = \begin{cases} 1 & (i, j \in E), \\ 0 & (i, j \notin E). \end{cases} \quad (7)$$

In the QUBO formulation, we search for the bit configuration $\{b_i\}$ that minimizes the cost function H . The cost function H is designed to decrease with the number of nodes selected (for maximizing the independent set) and increase if connected nodes with edges are selected (if the connected nodes are selected, the subset selected is no longer independent). The $\{b_i\}$ and H are defined as follows.

$$b_i = \begin{cases} 1 & (\text{Node } i \text{ is an element of the independent set}), \\ 0 & (\text{Node } i \text{ is not an element of the independent set}), \end{cases} \quad (8)$$

$$H = \sum_i^N \sum_j^N Q_{i,j} b_i b_j = H_A + H_B, \quad (9)$$

$$H_A = A \sum_i^N \sum_j^N f_{i,j} b_i b_j, \quad (10)$$

$$H_B = -B \sum_i^N b_i. \quad (11)$$

The QUBO cost function H is expressed with the coefficient matrix $\{Q_{i,j}\}$ and also the combination of H_A and H_B . The number of the constituents in the universe is N . Note that $b_i^2 = b_i$ for diagonal terms ($i = j$). H_A increases (gives a penalty) when a connected node is selected and is minimized when the subsets are formed only with unconnected nodes. H_B is proportional to the number of nodes selected; the higher the number, the lower the cost. A and B are positive coefficients and must be $B < A$ to avoid decreasing the cost function despite constraint violations (selecting connected nodes). In this work, $A = 2$ and $B = 1$.

3) STOCK UNIVERSE

In the following sections (Sec. V and Sec. VI), we use stock universes of various sizes corresponding to several market indices (TOPIX 100, Nikkei 225, TOPIX 1000, and TOPIX) for the Tokyo Stock Exchange (TSE). See APPENDIX A for the details on the market indices.

Since the constituents of market indices sometimes change, we determined the constituents of the stock universes as follows. The candidates of the stock universe for each market index are the constituents of the market index as of May 2023. From the candidates, the stocks that have remained listed for the analysis period consist of the stock universe. The size of the stock universe corresponding to TOPIX in Sec. V is 2,026, but that in Sec. VI is 1,747. This difference comes from the difference in the analysis periods; the analysis period in Sec. V is 2017 to 2023, while that in Sec. VI is 2010 to 2023.

C. PORTFOLIO SIMULATION

Figure 2 shows the flowchart for the simulation of the MIS portfolio strategy and the examples of data in the process steps (a) to (g) when assuming a 952-stock universe. In Sec. VI, in order to determine the best θ and asset-allocation method, we repeat the 10-year (120 months) backcast simulation while changing θ (19 patterns) for the constituents of TOPIX as the universe and then calculate the annualized return and risk from the 10-year simulation data.

At the beginning of the simulation, θ is set [the data (a)] and then the simulation evaluates the performance of the strategy every month. At the beginning of the monthly processing, the correlation matrix C [the data (b)] and the volatility $\{v_i\}$ [the data (c)] are calculated based on the price data for the last three years before the month as described in Sec. III-B1 and Sec. III-A2. The market graph [the data (d)] is constructed from the correlation matrix C by determining edge connections with the θ and then the Ising problem to be solved [the data (e)] is generated from the market graph (mapping of the MIS problem for the market graph onto the Ising problem) as described in Sec. III-B2. The SBM solves the Ising problem 10 times with different initial states and obtains 10 different solutions. Note that SB is a heuristic algorithm and can output different solutions by changing the initial states. After verifying the solutions in terms of constraint violation (if connected nodes are included), we select the best solution having the largest size

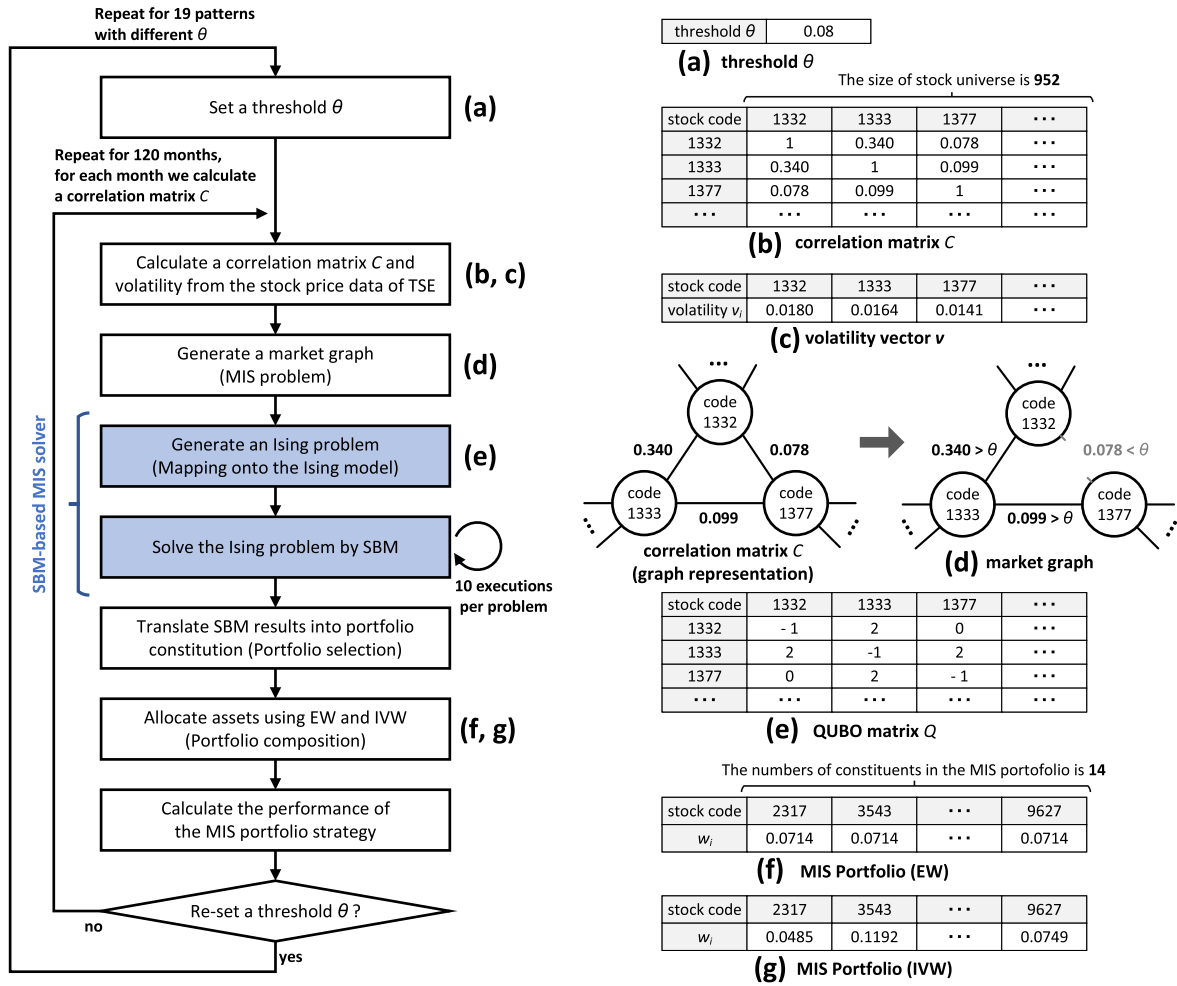


FIGURE 2. Flowchart for simulation of MIS portfolio strategy and the examples of data in the process steps (a) to (g). The data examples are for a 952-stock universe.

of the independent set from the verification-passed solutions. The best solution is translated directly into the portfolio constitution (portfolio selection) and then the compositions of the portfolio are determined for the two cases of the asset allocation methods (EW and IVW) [the data (f) and (g)]. Finally, we evaluate the new portfolio (considering the trading costs) and calculate the monthly return by comparing the evaluation value of the new portfolio with one of the last-month portfolio as described in Sec. III-A3.

IV. SYSTEM ARCHITECTURE

A. PORTFOLIO SIMULATOR

Figure 3 shows the architecture of the MIS portfolio simulator which executes the simulation process in the flowchart (Fig. 2) using the historical stock price data of the TSE. The simulator physically consists of a CPU-based server and an FPGA card installed in the PCIe slot of the server, which is logically modeled as three layers of software, middleware, and hardware. See APPENDIX B for the details of the hardware used.

The SBM (hardware layer) is implemented as a massively-parallel custom circuit with the FPGA and the server as

a look-aside accelerator (offloader) to the CPU-based host system. In the flowchart, the computationally-hard part, namely solving the NP-hard Ising problem (MIS problem), is offloaded to the SBM (dedicated accelerator) to reduce the overall simulation time. The remaining in the flowchart is executed by CPU processing (software layer). To solve the MIS problem in the monthly market graph by the SBM (an Ising machine), we have to map the MIS problem onto the Ising model. The mapping is executed also by the CPU. Thus, the SBM-based MIS solver (blue highlighted in Figs. 2 and 3) consists of mapping the MIS problem and solving the Ising problem, unlike the other software MIS solvers that directly solve the MIS problems. In the performance comparison of the MIS solvers (Sec. V), we include the times of mapping and communication needed for the SBM-based solver in the computation-time.

To make the SBM accessible at a higher abstraction level to the simulator program running on the CPU, we developed the middleware layer, namely, the hardware abstraction layer (HAL) for the SBM and the APIs (Application Programming Interface) for C/C++ and Python programming languages. Seen from the CPU side, the HAL abstracts the detailed

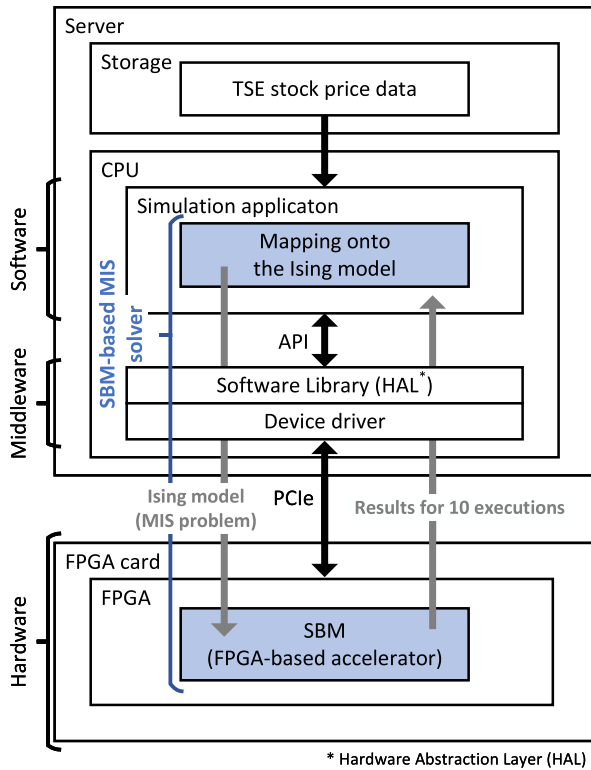


FIGURE 3. System architecture of the MIS portfolio simulator.

functionality of the device driver for the SBM. The APIs receive the SBM execution parameters such as the input data describing the Ising problem and the number of times to consecutively execute the SBM (10 times in this work), and then return the results (solutions) for the executions to the simulation program running in the software layer. More precisely, the API function calls the functions of the device driver to control PCIe communication between the CPU and the FPGA card and run the SBM core circuit. The SBM solves the Ising problem based on the execution parameters received via the PCIe communication with different initial states generated by the internal random number generator and then transmits the results (for 10 executions) back to the host system.

B. SBM CORE CIRCUIT

We describe the core circuit architecture for the ballistic SB algorithm (bSB) [17]. The overall circuit architecture for bSB follows the basic SBM design for adiabatic SB (aSB) [16], but the data path (pipeline path) for the time-evolution part is totally modified to represent the processing of the time-evolution producer specific to the bSB. In the physical representation of the bSB, a perfectly inelastic wall has been introduced to the mechanical system of the bSB. The circuit architecture for bSB has not been illustrated in the paper that introduced the bSB algorithm [17].

Simulated bifurcation [15], [17] numerically simulates the time-evolution of N nonlinear oscillators according to the Hamiltonian equations of motion, where the nonlinear oscillators correspond to the spin variables and the state of

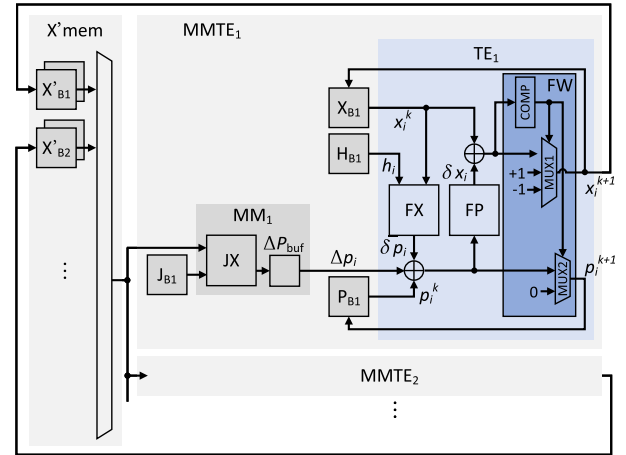


FIGURE 4. Core circuit architecture of the bSB accelerator.

i th oscillator is described by the position and momentum (x_i, p_i) . The SB time-evolution step consists of calculating the correction of momenta $\{\Delta p_i\}$ based on the many-body interaction [computationally corresponding to the matrix-vector multiplication (MM) of the J coupling matrix and $\{x_i\}$ position vector] and calculating the updated (time-evolved: TE) state variables, $\{x_i^{k+1}\}$ and $\{p_i^{k+1}\}$, from the $\{\Delta p_i\}$, bias $\{h_i\}$, and the current state variables, $\{x_i^k\}$ and $\{p_i^k\}$. After executing the time-evolution steps for a predetermined number (N_{step}), the positions of oscillators are digitized to be spins [the sign of x_i , $sgn(x_i)$, gives the state of i th spin].

Figure 4 shows the block diagram of the SB core circuit, where the pipeline path for the time-evolution part specific to the bSB is blue highlighted. The main two computation components are MM units corresponding to the multiply-accumulate (MAC) operations of $\sum_{j=1}^N J_{ij}x_j$ and TE pipelines corresponding to the time-evolution operation, which are combined to be MMTE units (each responsible for updating a subgroup of nonlinear oscillators). The MMTE units are organized with the global X'_{mem} memory unit to make a circulative structure as a whole corresponding to the iteration of the SB time-evolution steps. Memory modules (connected to TE) $X_B, P_B,$ and H_B store $x_i, p_i,$ and h_i , respectively.

The feature of bSB is the perfectly inelastic walls existing at the positions of $+1$ or -1 . If i th oscillator collides with the wall, the position and momentum of the oscillator are replaced by predetermined values $[x_i \leftarrow sgn(x_i), p_i \leftarrow 0]$. As shown in Fig. 4, the TE pipeline consists of the FX (calculates the value to update $p_i, \delta p_i$, from x_i), FP (calculates the value to update $x_i, \delta x_i$, from p_i), and FW (corresponds to the wall) modules, where the functionalities are expressed as follows.

$$FX(x_i, h_i) = \Delta t \{-(\alpha_0 - \alpha)x_i - \eta h_i\}, \tag{12}$$

$$FP(p_i) = \Delta t p_i, \tag{13}$$

$$FW(x_i) = \begin{cases} sgn(x_i) & (\text{when } |x_i| > 1), \\ x_i & (\text{when } |x_i| \leq 1), \end{cases} \tag{14}$$

$$FW(x_i, p_i) = \begin{cases} 0 & (\text{when } |x_i| > 1), \\ p_i & (\text{when } |x_i| \leq 1), \end{cases} \quad (15)$$

where Δt , α_0 , α and η are the SB parameters, the same as in the previous work [16]. The conditional-branch processing of the FW is implemented with one 32-bit comparator (COMP) and two 32-bit multiplexers (MUX1 and MUX2). The FX, FP, and FW modules are interconnected as shown with arrows in Fig. 4. The circuit inside the TE module corresponds to the hardwired algorithm of the bSB time-evolution operation.

The SBM implemented is capable of solving fully-connected 2,048-spin Ising problem with a computational precision of 32-bit floating-point, which works at the system clock frequency of 239 MHz in a highly-parallel fashion; the MM units have totally 4,096 MAC components (basic processing elements), which compute 4,096 spin-spin interactions in a single clock and there are 64 TE pipelines, including 64 sets of COMP, MUX1, and MUX2. The circuit resources used for the perfectly inelastic wall are not significant compared to the MAC components. The parameters for solving the MIS problem are $N_{step} = 1000$, $\Delta t = 0.2$ and $\eta = 0.2$.

V. MIS SOLVER PERFORMANCE

The SBM-based MIS solver is evaluated in terms of computation-time and solution-accuracy through comparison with conventional MIS solvers based on heuristic and exact algorithms (NetworkX [49] and OR-tools [50]).

NetworkX [49] is an open-source Python library for analyzing and visualizing graph structures, including a heuristic solver for the MIS problem. The size of an approximate solution for the MIS problem in a graph $G = (V, E)$ with nodes V and edges E is $O(|V|/(\log |V|)^2)$ in the worst case [52]. OR-tools [50] provides a solver for 0-1 integer programming problems applicable to the MIS problem, which is based on a branch-and-cut algorithm [53] known as an exact-solution method. Hereafter, NetworkX and OR-tools refer to the heuristic and exact-solution solvers for the MIS problem above mentioned. See APPENDIX B for the details of the execution.

Figure 5 shows the comparison results of the SBM-based MIS solver (heuristic), NetworkX (heuristic), and OR-tools (exact-solution) in terms of computation-time and solution-accuracy. The four sizes of stock universes corresponding to the market indices (TOPIX 100, Nikkei 225, TOPIX 1000, and TOPIX) were used and 10 market graphs (having the same size but different edges) for each universe were generated with $\theta = 0.25$ based on the correlation matrices for the different periods (consecutive three years, starting from different dates) in Mar. 2017 to May 2023. The numbers of nodes (corresponding to the number of the stocks in the universe) in the market graphs are 99, 222, 952, and 2,026 (see Sec. III-B3). The computation-time and the size of the independent set found (the larger, the better) when solving the MIS problem for each of the 10 market graphs were measured and averaged over the 10 market graphs. As mentioned in

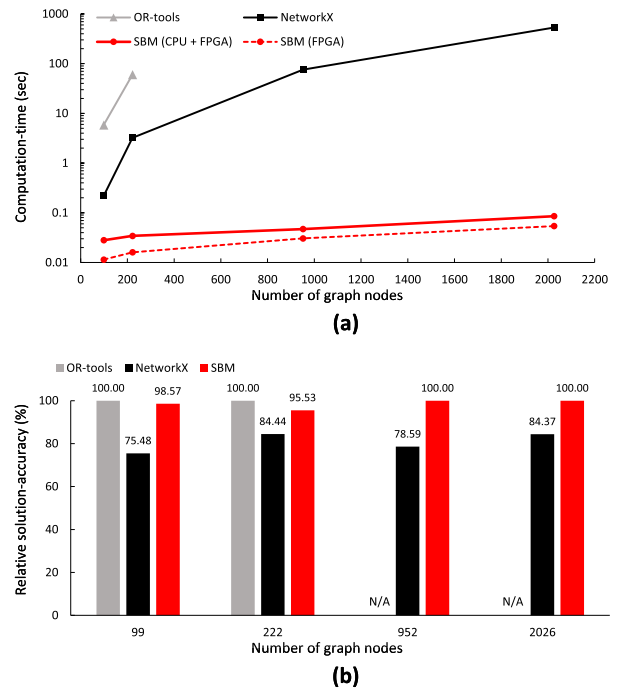


FIGURE 5. Performance comparison of the three MIS solvers, SBM (heuristic), NetworkX (heuristic), and OR-tools (exact-solution), in terms of (a) computation-time and (b) relative solution-accuracy (the size of the independent set found). The computation-time (the shorter, the better) and the size of the independent set found (the larger, the better) when solving the MIS problem for each of the 10 market graphs having the same size but different edges were measured and averaged over the 10 market graphs. Each value in (b) is the ratio to the largest one of the (averaged) sizes of independent sets found by the three solvers.

Sec. IV-A, the computation-time per market graph for the SBM [the red solid line in Fig. 5(a)] includes the time of 10 executions and the times of mapping and communication. The computation-time of SBM without the time of mapping is also shown there as the red dashed line. The total number of SBM executions for Fig. 5 is 400 (10 executions per graph, 10 market graphs per universe, 4 universes).

From the computation-times in Fig. 5 (a), it is observed that SBM is the fastest for all the sizes of market graphs. OR-tools could not output solutions for problems with 952 nodes or more by a timeout error (we define the timeout time as 10 hours). The speed advantage of SBM becomes more pronounced as the number of nodes increases, and at 2,026 nodes, the average computation-times for SBM and NetworkX are 8.54×10^{-2} and 5.32×10^2 seconds, respectively, with SBM being 6,230 times faster. This difference is further pronounced in practical use cases because the performance of a portfolio strategy should be analyzed for a long-term period under various conditions for optimizing strategy parameters (as we will see an example in Sec. VI).

The solution-accuracy (the size of the independent set found) in Fig. 5 (b) is shown as the ratio to the largest one of the sizes of independent sets found by the three solvers. Since OR-tools is an exact-solution solver, the relative accuracy is 100% at the node numbers 99 and 222. At the 99 and 222 nodes, the relative accuracies of

TABLE 1. Statistics of the simulated 10-year (120-month) performance data of the MIS portfolio strategy when varying θ from 0.18 to 0.36 (19 patterns) for the two cases of the asset-allocation methods (EW or IVW), including the edge density (max, min, and average) of market graphs, the size (# of stock) of independent set found (max, min, average, and standard deviation: SD), the annualized return/risk and Sharpe ratio of the strategy.

θ	0.18	0.19	0.20	0.21	0.22	0.23	0.24	0.25	0.26	0.27	0.28	0.29	0.30	0.31	0.32	0.33	0.34	0.35	0.36
Edge density																			
Max (%)	94.9	93.6	92.0	90.2	88.1	85.7	83.0	80.1	76.9	74.1	71.2	68.2	65.0	61.8	58.6	55.2	51.8	48.4	44.9
Min (%)	71.7	68.2	64.7	61.0	57.2	53.4	49.6	45.8	42.0	38.4	34.9	31.6	28.4	25.5	22.7	20.2	17.8	15.7	13.8
Avr (%)	87.5	85.4	83.2	80.6	77.9	75.0	71.9	68.6	65.2	61.7	58.1	54.5	50.8	47.1	43.5	39.9	36.5	33.1	29.9
Stock No.																			
Max	132	151	173	196	225	248	278	312	350	387	420	454	501	539	593	640	680	738	784
Min	3	27	32	37	43	52	60	71	81	96	111	129	150	166	191	218	246	276	305
Avr	62	73	84	95	108	123	140	158	178	201	224	251	279	310	343	378	414	453	492
SD	35	37	42	47	52	58	64	70	77	84	91	97	103	110	117	123	129	135	140
Result (EW)																			
Return (%)	18.1	18.7	16.5	16.4	17.4	16.4	16.7	16.4	15.4	16.4	15.1	15.2	15.4	15.7	14.8	14.9	15.0	15.2	15.1
Risk (%)	17.1	17.0	16.3	16.5	16.0	15.7	15.9	15.8	15.8	15.5	15.6	15.4	15.3	15.4	15.3	15.3	15.3	15.2	15.2
Sharpe ratio	1.06	1.10	1.01	1.00	1.09	1.04	1.05	1.04	0.98	1.05	0.97	0.98	1.00	1.02	0.96	0.97	0.99	1.00	0.99
Result (IVW)																			
Return (%)	19.6	18.2	16.8	16.6	17.2	16.3	16.5	16.2	15.1	15.9	14.6	14.7	14.8	14.9	14.0	14.1	14.2	14.4	14.2
Risk (%)	15.5	15.0	14.2	14.7	14.2	14.0	14.2	14.0	14.1	13.8	13.9	13.8	13.8	13.9	13.8	13.9	13.8	13.8	13.8
Sharpe ratio	1.26	1.21	1.18	1.13	1.21	1.16	1.16	1.16	1.07	1.15	1.05	1.06	1.07	1.08	1.01	1.02	1.03	1.04	1.03

the SBM are close to those of OR-tools, while NetworkX is considerably inferior to the other two solvers. More specifically, at the 99 and 222 nodes, the SBM found the exact solutions (the same solutions as ones by OR-tools) for 9 and 6 samples (out of 10 samples for each), respectively, and the sizes of the remaining solutions (1 and 4 samples for the 99 and 222 nodes, respectively) are smaller only by 1 from the sizes of the exact solutions; the SBM provide approximate solutions very close to the exact solutions. The accuracy advantage of SBM over NetworkX seen at the 99 and 222 nodes (roughly 20%) is maintained almost constant for larger problem sizes of the 952 and 2026 nodes (OR-tools was not capable of solving those sizes of problems).

The independent sets in undirected market graphs are considered in this work, while clique-like structures in directed graphs (sometimes called dicliques) have been investigated [54], [55] [independent set problems are in a complementary relationship with clique problems]. Market representation as directed graphs and the selection of uncorrelated stocks in those directed graphs would be one of the interesting future works.

VI. PORTFOLIO PERFORMANCE

By using the SBM-based portfolio simulator described in Sec. IV-A, we simulate the performance of the MIS portfolio strategy over 10 years (Apr. 1, 2013, to Mar. 31, 2023) for a large universe including 1,747 stocks corresponding to TOPIX. The strategy parameter (correlation threshold θ) and the asset-allocation methods (EW or IVW) are optimized by repeating the long-term, large-scale backcast simulation. Then we compare the performance of the best MIS portfolio strategy with the Japanese major indices (assuming passive index funds) from the perspective of risk-return characteristics.

A. PARAMETER SEARCH

Table 1 shows the statistics of the simulated 10-year (120-month) performance data of the MIS portfolio strategy when varying θ from 0.18 to 0.36 (19 patterns) for the two cases of the asset-allocation methods (EW or IVW) [totally 38 settings]. As illustrated in Fig. 2, the total number of SBM executions for Table 1 is 22,800 (10 executions per graph, 120 market graphs per θ , 19 patterns of θ). For each of the 120 market graphs, the edge density and the size (the number of stocks) of the independent set found were examined and their statistics are summarized in Table 1. As the correlation threshold θ increases, the number of excluded edges increases, and therefore the edges get more sparse, as a result, the size of the independent set increases. These trends are consistent with the previous studies [6], [7], [8].

Figure 6 shows the cumulative monthly returns of the MIS portfolio strategies for the representative six settings as a function of time. The performance of the strategy, i.e., the annualized return/risk and Sharpe ratio, are calculated from the data of the monthly returns, where the Sharpe ratio [56] is, in this work, the ratio of the mean to the standard deviation of the return (the profit and loss per period for an investment) from a strategy as in [57]. For both EW and IVW, the risks are reduced as θ increases. This can be interpreted as an increase in constituents of the portfolio leading to a reduction in risk, which is the diversification effect as expected. The Sharpe ratio, as well as the risk, peaks at $\theta = 0.18$ or 0.19 and then slightly declines as θ increases, although it remains above 1 for most of the θ settings.

In this work, we select the combination of $\theta = 0.23$ and IVW as one that gives the best performance (Sharpe ratio = 1.16) under a constraint that the numbers of constituents in the portfolios should be more than 50 for the diversification effect. When comparing the EW and IVW methods, the annualized returns are comparable, while the annualized

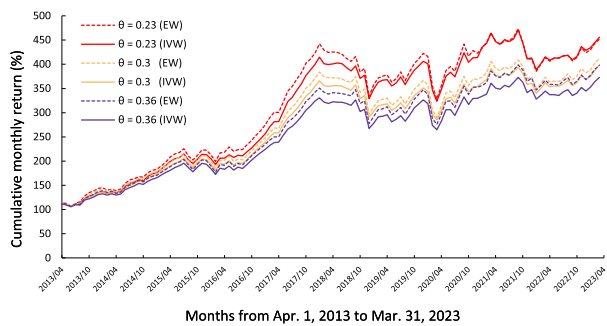


FIGURE 6. Cumulative monthly returns of the MIS portfolio strategies with one of the two asset-allocation methods (IVW and EW) when varying the correlation threshold θ . Simulation data is from Apr. 1, 2013, to Mar. 31, 2023 (10 years).

risks are, overall, better for the IVW method. The results demonstrate the validity of the IVW concept of predicting future risks based on past volatility.

B. PERFORMANCE EVALUATION

Figure 7 shows the cumulative monthly returns of the best MIS portfolio strategy ($\theta = 0.23$ and IVW) and major indices in the TSE [TOPIX, TOPIX 1000, TOPIX 500, TOPIX Mid400, TOPIX Small, and MSCI Japan Minimum Volatility Index (see APPENDIX A for the details of the indices)]. Their risk-return characteristics (annualized return/risk and Sharpe ratio) are summarized in Table 2 and Fig. 8. The comparison in Table 2 and Fig. 8 is conservative for the MIS strategy because the trading cost is included only in the MIS strategy (see Sec. III-A3) [note that the dividends are considered for the MIS strategy and the major indices].

The results clearly show that the MIS strategy outperforms all the indices in terms of annualized return and Sharpe ratio. The MIS strategy is superior to the TOPIX series indices because of the relatively high return and relatively low risk. It is remarkable that the MIS strategy (a diversified portfolio strategy) outperforms the risk-oriented index, MSCI Japan Minimum Volatility Index, in terms of the Sharpe ratio due to the relatively high return.

We analyzed the factors contributing to the relatively high return of the MIS strategy. Figure 9 shows the cumulative excess return for the MIS strategy versus TOPIX. The annualized return of the MIS strategy for the period from Sep. 2016 and Dec. 2017 is 49.0%, remarkably larger than that of TOPIX (22.7%).

To examine which and what kind of stocks contributed to the return of the MIS strategy, we introduce an index, *DIFR*, for each stock. *DIFR*_{*i*} is defined as the difference in the return of stock *i* between the MIS portfolio strategy and TOPIX for the period and expressed by

$$DIFR_i = R_i^{MIS} - R_i^{TPX}, \tag{16}$$

$$R_i^{MIS} = \sum_t R_i(t) \cdot w_i^{MIS}(t), \tag{17}$$

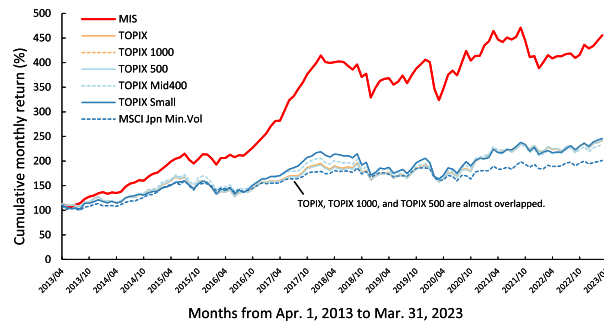


FIGURE 7. Cumulative monthly returns of the best MIS portfolio strategy ($\theta = 0.23$ and IVW) and major indices in the TSE (TOPIX, TOPIX 1000, TOPIX 500, TOPIX Mid400, TOPIX Small, and MSCI Japan Minimum Volatility Index). The evaluation is from Apr.1, 2013, to Mar. 31, 2023 (10 years).

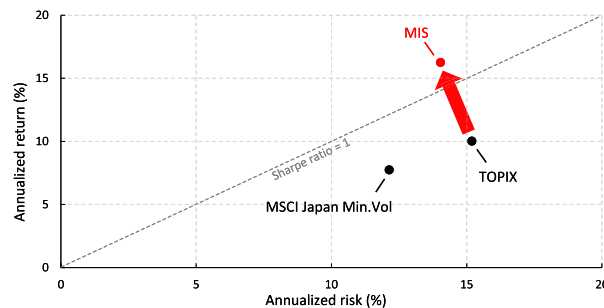


FIGURE 8. Mapping of the best MIS portfolio strategy and major TSE indices (the data in Table 2) on the risk-return graph. Since the data for TOPIX 1000, TOPIX 500, TOPIX Mid400, and TOPIX Small are almost overlapped with that of TOPIX, the data for TOPIX is only illustrated.

$$R_i^{TPX} = \sum_t R_i(t) \cdot w_i^{TPX}(t). \tag{18}$$

Here, $R_i(t)$ is the monthly return of stock *i* at month *t* (*T* is 16 months in the period). $w_i^{MIS}(t)$ and $w_i^{TPX}(t)$ are the weights of stock *i* in the MIS portfolio and TOPIX at month *t*, respectively. R_i^{MIS} and R_i^{TPX} are the returns of stock *i* in the MIS strategy and TOPIX (a TOPIX-associated passive fund), respectively.

The size of the universe for the MIS strategy is 1,747 (see Sec. III-B3), while the constituents of TOPIX is 2,156 as of Jul. 2023 (see APPENDIX A). Assuming the constituents of TOPIX are the same as the universe of the MIS strategy, we calculated the performance and weights [$w_i^{TPX}(t)$] for TOPIX according to the weighting method of TOPIX (a capitalization-based weighting, see APPENDIX A). Hence, $\sum_{i=1}^{1747} w_i^{MIS}(t) = 1$ and $\sum_{i=1}^{1747} w_i^{TPX}(t) = 1$. We have confirmed that the performance of TOPIX for the 1,747 constituents is almost the same as that of the original TOPIX since the dominant constituents do not differ.

Table 3 shows the result of a differential factor analysis of the performances between the MIS portfolio strategy and TOPIX for the period. The top five and bottom five stocks when sorted by *DIFR* are listed there, along with the market capitalization, the degree of the node in the market graph, w^{MIS} , and w^{TPX} (averaged over the period). Here, the degree of a node, in graph theory, represents the number of edges that are connected to the node.

TABLE 2. Annualized return, risk, and Sharpe ratio of the best MIS portfolio strategy ($\theta = 0.23$ and IVW) and major indices in the TSE (TOPIX, TOPIX 1000, TOPIX 500, TOPIX Mid400, TOPIX Small, and MSCI Japan Minimum Volatility Index). The evaluation is from Apr.1, 2013, to Mar. 31, 2023 (10 years).

	TOPIX	MSCI Jpn Min.Vol	TOPIX 500	TOPIX 1000	TOPIX Mid400	TOPIX Small	MIS Portfolio
Return (%)	10.0	7.7	10.0	10.0	9.7	10.2	16.3
Risk (%)	15.2	12.1	15.3	15.3	15.2	15.2	14.0
Sharpe ratio	0.66	0.64	0.65	0.66	0.64	0.67	1.16

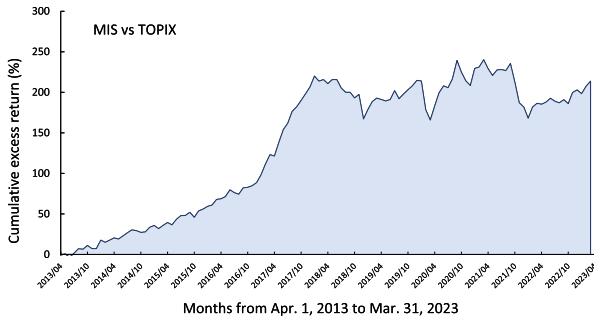


FIGURE 9. The cumulative excess return of MIS portfolio versus TOPIX. The evaluation is from Apr.1, 2013, to Mar. 31, 2023 (10 years).

The top five stocks are all small-capitalization stocks and the bottom five are all large-capitalization stocks. We have confirmed that the top five small-capitalization stocks are involved in the MIS strategy but the bottom five large-capitalization stocks are not in the MIS strategy (also see w^{MIS} in Table 3). As partially shown in Table 3 (market capitalization and the degree of the node), it has been observed that the degree of the node tends to increase with the capitalization of the stock increases. It would be because the large-capitalization stocks have large influences on the market and are correlated with one another (associated strongly with the whole market). The stocks with the smaller degrees of the node are more likely to be selected for the MIS portfolio. Hence, the MIS strategy is more likely to be composed of small-capitalization and low-correlation stocks. In contrast, TOPIX is composed of large-capitalization stocks due to the definition of the weighting method [TOPIX Small is composed of small-capitalization stocks (not considering correlations)]. It is found that, in the period, the selection of small-capitalization and low-correlation stocks results in the superior return of the MIS strategy seen in Figs. 7 and 8. One of the interesting future works would be to investigate whether this phenomenon (the MIS strategy may lead to not only a low risk but also a high return) also occurs in other stock markets (such as the New York Stock Exchange and the London Stock Exchange, etc.) and other kinds of markets for various financial products.

VII. CONCLUSION

We have developed a look-aside type combinatorial optimization accelerator (an Ising machine) based on a quantum-inspired parallelizable algorithm called simulated bifurcation (SB), which enables the long-term backcast simulation of a diversified portfolio strategy based on the selection of the maximum independent set in a large-scale market

graph representing the correlations between stocks (the MIS problem is known to be NP-hard). We optimized the strategy parameters of the MIS portfolio strategy by iterating the long-term large-scale simulations and found that the best MIS strategy outperforms the major market indices such as TOPIX, TOPIX Small, and MSCI Japan Minimum Volatility Index.

The look-aside type accelerator has a massively-parallel custom circuit (core circuit) for the ballistic SB (a variant of the SB algorithms) featuring the time-evolution pipeline path specific to the ballistic SB (corresponding to a perfectly inelastic wall in physics) and has been implemented as a PCIe-attachable FPGA card. We also have developed the hardware abstraction layer (HAL) and the APIs (Application Programming Interface), making the SB-based FPGA accelerator accessible at a higher abstraction level to the financial simulation program running on the CPU.

We systematically evaluated the performance of the SB-based Ising machine (fully-connection 2,048-spin size, 32-bit floating point precision) as a MIS solver using various sizes of practical market graphs (generated from the historical market data). The SB-based MIS solver provides a good approximate solution very close to the exact solution in a much shorter time than an exact-solution solver (OR-tools) and also finds the independent sets with a remarkably larger size (roughly 20%) than a conventional heuristic solver (NetworkX). At a large-scale market graph of 2,026 nodes (too large for the exact solver to handle), the SB-based solver is 6,230 times faster than the conventional heuristic solver.

By using the SB-based solver, we investigated the performance of the MIS portfolio strategy for a large-scale universe including 1,747 stocks (corresponding to the constituents of a major market index, TOPIX) by repeating the 10-year simulation (2013 to 2023) while varying correlation threshold parameters (19 patterns) and the asset-allocation methods (inverse volatility weight or equal weight) [totally involving the 22,800 executions of the SB solver]. By comparing with the major market indices, it has been found that the best MIS portfolio strategy based on the large-scale universe has not only a relatively low risk (the diversification effect as expected) but also a relatively high return, outperforming the major indices. We analyzed the factors contributing to the relatively high return of the MIS strategy and concluded that the selection of small-capitalization and low-correlation stocks results in the superior return.

A potential direction of further research is to examine the performance and risk-return characteristics of the MIS portfolio strategy for a large-scale universe in other stock markets

TABLE 3. Deferential factor analysis of the performances between the MIS portfolio strategy and TOPIX. The difference in the return of a stock between the MIS portfolio strategy and TOPIX for a period from Sep. 2016 to Dec. 2017 is defined as *DIFR* (see the main text). Listed are company name, stock code, market capitalization (in Japanese yen, JYN), the degree of the node in the market graph, weight in the MIS portfolio (w^{MIS}), weight in TOPIX (w^{TPX}) and *DIFR* for the top five and bottom five stocks when sorted by *DIFR*. The market capitalization, the degree of the node, w^{MIS} , and w^{TPX} are averaged over the period.

Rank	Company	Stock code	Market capitalization (Millions of JYN)	Degree of the node	w^{MIS} (%)	w^{TPX} (%)	<i>DIFR</i> (%)
1	IK Holdings Co Ltd	2722	8,523	340	0.511	0.002	+ 2.05
2	Imuraya Group Co Ltd	2209	15,774	0	1.850	0.005	+ 1.91
3	Kawanishi Warehouse Co Ltd	9322	5,681	1	1.159	0.001	+ 1.71
4	Jutec Holdings Corp	3157	5,391	12	1.081	0.002	+ 1.47
5	Seed Co Ltd	7743	5,701	4	0.909	0.001	+ 1.46
1743	Sony Group Corp	6758	4,472,465	1591	0.000	1.299	- 0.57
1744	Nintendo Co Ltd	7974	3,235,738	1325	0.000	0.921	- 0.57
1745	Keyence Corp	6861	3,673,789	1636	0.000	1.042	- 0.59
1746	Toyota Motor Corp	7203	12,625,742	1662	0.000	3.757	- 0.74
1747	Mitsubishi UFJ Financial Group Inc	8306	8,357,252	1591	0.000	2.453	- 1.04

TABLE 4. The constituents and weighting methods of the major market indices in the TSE. The numbers of component stocks are ones as of Jul. 2023.

Index	Number of Stocks	Component Stocks	Weighting Method
TOPIX	2,156	All companies in the prime market division of the TSE.	Weighted by free-float adjusted market capitalization which is formulated by “total number of shares - number of shares not available for trading by public * stock price”.
TOPIX 500	498	Top 500 stocks (large and mid-cap stocks) in TOPIX by market capitalization and liquidity.	Weighted by free-float adjusted market capitalization.
TOPIX 1000	993	Top 1,000 stocks in TOPIX by market capitalization and liquidity.	Weighted by free-float adjusted market capitalization.
TOPIX Mid400	398	400 stocks in the TOPIX 500, excluding the top 100 (large-cap stocks). The component stocks of the TOPIX Mid400 are called mid-cap stocks.	Weighted by free-float adjusted market capitalization.
TOPIX Small	1,658	Stocks from TOPIX excluding TOPIX 500. The component stocks of the TOPIX Small are called small-cap stocks.	Weighted by free-float adjusted market capitalization.
MSCI Jpn Min.Vol	139	Selected from large and mid-cap stocks in the TSE according to portfolio strategy.	Weighted by using the Barra Global Equity Model [58], for the lowest absolute risk.
Nikkei 225	225	225 stocks with high liquidity in the prime market division of the TSE.	Weighted by stock price which is adjusted by the price adjustment factor.

or other kinds of markets for various financial products. Representing markets as more complicated graphs/networks like directed graphs or higher-order networks would be another direction of further research.

APPENDIX A MARKET INDICES IN THE TSE

Table 4 summarizes the constituents and weighting (composition) of the major market indices in the TSE (TOPIX, TOPIX 500, TOPIX 1000, TOPIX Mid400, TOPIX Small, MSCI Japan Minimum Volatility Index, and Nikkei 225). The numbers of the component stocks are ones as of Jul. 2023.

APPENDIX B IMPLEMENTATION DETAILS

The SBM-based MIS solver consists of mapping onto an Ising model by software processing (CPU) and solving the Ising problem corresponding to the MIS problem by hardware processing (FPGA). The FPGA (Intel Stratix 10 SX 2800 FPGA) on the board (Intel FPGA PAC D5005 accelerator card) has 933,120 adaptive logic modules (ALMs) including

2,753,000 adaptive look-up-tables (ALUTs, 5-input LUT equivalent) and 3,732,480 flip-flop registers, 11,721 20Kbit-size RAM blocks (BRAMs), and 5,760 digital signal processor blocks (DSPs). The SBM components in the FPGA described in Sec. IV-B were coded in a high-level synthesis (HLS) language (Intel FPGA SDK for OpenCL, ver. 19.2). The PCIe configuration of the FPGA card is PCIe Gen3 \times 16 with a peak bandwidth of 15.75 GB/s. The software processing is executed with a single CPU [Intel Core i3-10100 (3.60 GHz, 4 cores)] and 8 GB DDR-DRAM.

The software MIS solvers of NetworkX (ver. 2.5.1) and OR-tools (ver. 9.4.1874) implemented as the Python libraries are executed on a server with dual CPUs [Intel Xeon Silver 4215R (3.20 GHz, 8 cores)] and 256 GB DDR-DRAM.

ACKNOWLEDGMENT

The authors would like to thank Masaya Yamasaki and Hayato Goto for their helpful comments.

CONFLICTS OF INTEREST

R.H., Y.H., and K.T. are included as inventors on two U.S. patent applications related to this work filed by the Toshiba

Corporation (no. 17/249353, filed 20 February 2020; no. 17/249293, filed 25 February 2021). The authors declare that they have no other competing interests.

REFERENCES

- [1] H. Markowitz, "Portfolio selection," *J. Finance*, vol. 7, no. 1, pp. 77–91, Mar. 1952, doi: [10.1111/j.1540-6261.1952.tb01525.x](https://doi.org/10.1111/j.1540-6261.1952.tb01525.x).
- [2] W. F. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *J. Finance*, vol. 19, no. 3, pp. 425–442, Sep. 1964, doi: [10.1111/j.1540-6261.1964.tb02865.x](https://doi.org/10.1111/j.1540-6261.1964.tb02865.x).
- [3] G. B. Koumou, "Diversification and portfolio theory: A review," *Financial Markets Portfolio Manage.*, vol. 34, no. 3, pp. 267–312, Sep. 2020, doi: [10.1007/s11408-020-00352-6](https://doi.org/10.1007/s11408-020-00352-6).
- [4] Y.-H. Chou, S.-Y. Kuo, and Y.-T. Lo, "Portfolio optimization based on funds standardization and genetic algorithm," *IEEE Access*, vol. 5, pp. 21885–21900, 2017, doi: [10.1109/ACCESS.2017.2756842](https://doi.org/10.1109/ACCESS.2017.2756842).
- [5] B. Jin, "A mean-VaR based deep reinforcement learning framework for practical algorithmic trading," *IEEE Access*, vol. 11, pp. 28920–28933, 2023, doi: [10.1109/ACCESS.2023.3259108](https://doi.org/10.1109/ACCESS.2023.3259108).
- [6] S. Butenko, "Maximum independent set and related problems, with applications," Ph.D. dissertation, Dept. Ind. Syst. Eng., Univ. Florida, Gainesville, FL, USA, 2003. [Online]. Available: <https://dl.acm.org/doi/10.5555/1023044>
- [7] V. Boginski, S. Butenko, and P. M. Pardalos, "Network-based techniques in the analysis of the stock market," in *Supply Chain and Finance*. Singapore: World Scientific, 2004, pp. 1–14, doi: [10.1142/9789812562586_0001](https://doi.org/10.1142/9789812562586_0001).
- [8] V. Boginski, S. Butenko, and P. M. Pardalos, "Statistical analysis of financial networks," *Comput. Statist. Data Anal.*, vol. 48, no. 2, pp. 431–443, Feb. 2005, doi: [10.1016/j.csda.2004.02.004](https://doi.org/10.1016/j.csda.2004.02.004).
- [9] V. Kalyagin, A. Koldanov, P. Koldanov, and V. Zamaraev, "Market graph and Markowitz model," in *Optimization in Science and Engineering: In Honor of the 60th Birthday of Panos M. Pardalos*. New York, NY, USA: Springer, 2014, pp. 293–306, doi: [10.1007/978-1-4939-0808-0_15](https://doi.org/10.1007/978-1-4939-0808-0_15).
- [10] M. Marzec, "Portfolio optimization: Applications in quantum computing," in *Handbook of High-Frequency Trading and Modeling in Finance*. NJ, USA: Wiley, 2016, pp. 73–106, doi: [10.1002/9781118593486.ch4](https://doi.org/10.1002/9781118593486.ch4).
- [11] S. Butenko, P. Pardalos, I. Sergienko, V. Shylo, and P. Stetsyuk, "Finding maximum independent sets in graphs arising from coding theory," in *Proc. ACM Symp. Appl. Comput.*, Mar. 2002, pp. 542–546, doi: [10.1145/508791.508897](https://doi.org/10.1145/508791.508897).
- [12] K.-Y. Lee and T.-C. Wang, "Post-routing redundant via insertion for yield/reliability improvement," in *Proc. Asia South Pacific Conf. Design Autom. Conf.*, 2006, pp. 303–308, doi: [10.1109/ASPDAC.2006.1594699](https://doi.org/10.1109/ASPDAC.2006.1594699).
- [13] J. Zhou, L. Wang, W. Wang, and Q. Zhou, "Efficient graph-based resource allocation scheme using maximal independent set for Randomly-deployed small star networks," *Sensors*, vol. 17, no. 11, p. 2553, Nov. 2017, doi: [10.3390/s17112553](https://doi.org/10.3390/s17112553).
- [14] A. Mallick, M. K. Bashar, D. S. Truesdell, B. H. Calhoun, S. Joshi, and N. Shukla, "Using synchronized oscillators to compute the maximum independent set," *Nature Commun.*, vol. 11, no. 1, p. 4689, Sep. 2020, doi: [10.1038/s41467-020-18445-1](https://doi.org/10.1038/s41467-020-18445-1).
- [15] H. Goto, K. Tatsumura, and A. R. Dixon, "Combinatorial optimization by simulating adiabatic bifurcations in nonlinear Hamiltonian systems," *Sci. Adv.*, vol. 5, no. 4, p. eaav2372, Apr. 2019, doi: [10.1126/sciadv.aav2372](https://doi.org/10.1126/sciadv.aav2372).
- [16] K. Tatsumura, A. R. Dixon, and H. Goto, "FPGA-based simulated bifurcation machine," in *Proc. 29th Int. Conf. Field Program. Log. Appl. (FPL)*, Sep. 2019, pp. 59–66, doi: [10.1109/FPL.2019.00019](https://doi.org/10.1109/FPL.2019.00019).
- [17] H. Goto, K. Endo, M. Suzuki, Y. Sakai, T. Kanao, Y. Hamakawa, R. Hidaka, M. Yamasaki, and K. Tatsumura, "High-performance combinatorial optimization based on classical mechanics," *Sci. Adv.*, vol. 7, no. 6, p. eabe7953, Feb. 2021, doi: [10.1126/sciadv.abe7953](https://doi.org/10.1126/sciadv.abe7953).
- [18] K. Tatsumura, M. Yamasaki, and H. Goto, "Scaling out ising machines using a multi-chip architecture for simulated bifurcation," *Nature Electron.*, vol. 4, no. 3, pp. 208–217, Mar. 2021, doi: [10.1038/s41928-021-00546-4](https://doi.org/10.1038/s41928-021-00546-4).
- [19] T. Kanao and H. Goto, "Simulated bifurcation assisted by thermal fluctuation," *Commun. Phys.*, vol. 5, no. 1, p. 153, Jun. 2022, doi: [10.1038/s42005-022-00929-9](https://doi.org/10.1038/s42005-022-00929-9).
- [20] M. W. Johnson et al., "Quantum annealing with manufactured spins," *Nature*, vol. 473, no. 7346, pp. 194–198, May 2011, doi: [10.1038/nature10012](https://doi.org/10.1038/nature10012).
- [21] A. D. King et al., "Quantum critical dynamics in a 5,000-qubit programmable spin glass," *Nature*, vol. 617, no. 7959, pp. 61–66, May 2023, doi: [10.1038/s41586-023-05867-2](https://doi.org/10.1038/s41586-023-05867-2).
- [22] T. Honjo, T. Sonobe, K. Inaba, T. Inagaki, T. Ikuta, Y. Yamada, T. Kazama, K. Enbutsu, T. Umeki, R. Kasahara, K. Kawarabayashi, and H. Takesue, "100,000-spin coherent ising machine," *Sci. Adv.*, vol. 7, no. 40, p. eabh0952, Oct. 2021, doi: [10.1126/sciadv.abh0952](https://doi.org/10.1126/sciadv.abh0952).
- [23] D. Pierangeli, G. Marcucci, and C. Conti, "Large-scale photonic ising machine by spatial light modulation," *Phys. Rev. Lett.*, vol. 122, no. 21, May 2019, Art. no. 213902, doi: [10.1103/PhysRevLett.122.213902](https://doi.org/10.1103/PhysRevLett.122.213902).
- [24] F. Cai, S. Kumar, T. Van Vaerenbergh, X. Sheng, R. Liu, C. Li, Z. Liu, M. Foltin, S. Yu, Q. Xia, J. J. Yang, R. Beausoleil, W. D. Lu, and J. P. Strachan, "Power-efficient combinatorial optimization using intrinsic noise in memristor Hopfield neural networks," *Nature Electron.*, vol. 3, no. 7, pp. 409–418, Jul. 2020, doi: [10.1038/s41928-020-0436-6](https://doi.org/10.1038/s41928-020-0436-6).
- [25] N. A. Aadit, A. Grimaldi, M. Carpentieri, L. Theogarajan, J. M. Martinis, G. Finocchio, and K. Y. Camsari, "Massively parallel probabilistic computing with sparse ising machines," *Nature Electron.*, vol. 5, no. 7, pp. 460–468, Jun. 2022, doi: [10.1038/s41928-022-00774-2](https://doi.org/10.1038/s41928-022-00774-2).
- [26] W. Moy, I. Ahmed, P.-W. Chiu, J. Moy, S. S. Sapatnekar, and C. H. Kim, "A 1,968-node coupled ring oscillator circuit for combinatorial optimization problem solving," *Nature Electron.*, vol. 5, no. 5, pp. 310–317, May 2022, doi: [10.1038/s41928-022-00749-3](https://doi.org/10.1038/s41928-022-00749-3).
- [27] A. Sharma, R. Afoakwa, Z. Ignjatovic, and M. Huang, "Increasing ising machine capacity with multi-chip architectures," in *Proc. 49th Annu. Int. Symp. Comput. Archit.*, Jun. 2022, pp. 508–521, doi: [10.1145/3470496.3527414](https://doi.org/10.1145/3470496.3527414).
- [28] T. Takemoto, M. Hayashi, C. Yoshimura, and M. Yamaoka, "A 2 × 30k-spin multi-chip scalable CMOS annealing processor based on a Processing-in-Memory approach for solving large-scale combinatorial optimization problems," *IEEE J. Solid-State Circuits*, vol. 55, no. 1, pp. 145–156, Jan. 2020, doi: [10.1109/JSSC.2019.2949230](https://doi.org/10.1109/JSSC.2019.2949230).
- [29] K. Kawamura, J. Yu, D. Okonogi, S. Jimbo, G. Inoue, A. Hyodo, Á. L. García-Arias, K. Ando, B. H. Fukushima-Kimura, R. Yasudo, T. Van Chu, and M. Motomura, "Amorphica: 4-Replica 512 fully connected spin 336MHz metamorphic annealer with programmable optimization strategy and Compressed-Spin-Transfer multi-chip extension," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2023, pp. 42–44, doi: [10.1109/ISSCC42615.2023.10067504](https://doi.org/10.1109/ISSCC42615.2023.10067504).
- [30] K. Yamamoto, K. Kawamura, K. Ando, N. Mertig, T. Takemoto, M. Yamaoka, H. Teramoto, A. Sakai, S. Takamaeda-Yamazaki, and M. Motomura, "Statica: A 512-spin 0.25m-weight annealing processor with an all-spin-updates-at-once architecture for combinatorial optimization with complete spin-spin interactions," *IEEE J. Solid-State Circuits*, vol. 56, no. 1, pp. 165–178, Jan. 2021, doi: [10.1109/JSSC.2020.3027702](https://doi.org/10.1109/JSSC.2020.3027702).
- [31] S. Matsubara, M. Takatsu, T. Miyazawa, T. Shibasaki, Y. Watanabe, K. Takemoto, and H. Tamura, "Digital annealer for high-speed solving of combinatorial optimization problems and its applications," in *Proc. 25th Asia South Pacific Design Autom. Conf. (ASP-DAC)*, Jan. 2020, pp. 667–672, doi: [10.1109/ASP-DAC47756.2020.9045100](https://doi.org/10.1109/ASP-DAC47756.2020.9045100).
- [32] H. M. Waidyasoorya and M. Hariyama, "Highly-parallel FPGA accelerator for simulated quantum annealing," *IEEE Trans. Emerg. Topics Comput.*, vol. 9, no. 4, pp. 2019–2029, Oct. 2021, doi: [10.1109/TETC.2019.2957177](https://doi.org/10.1109/TETC.2019.2957177).
- [33] T. Okuyama, T. Sonobe, K. Kawarabayashi, and M. Yamaoka, "Binary optimization by momentum annealing," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 100, no. 1, Jul. 2019, Art. no. 012111, doi: [10.1103/PhysRevE.100.012111](https://doi.org/10.1103/PhysRevE.100.012111).
- [34] A. Lucas, "Ising formulations of many NP problems," *Frontiers Phys.*, vol. 2, p. 5, Feb. 2014, doi: [10.3389/fphy.2014.00005](https://doi.org/10.3389/fphy.2014.00005).
- [35] Y. Shang, "Characterization of expansion-related properties of modular graphs," *Discrete Appl. Math.*, vol. 338, pp. 135–144, Oct. 2023, doi: [10.1016/j.dam.2023.06.002](https://doi.org/10.1016/j.dam.2023.06.002).
- [36] T. Kanao and H. Goto, "Simulated bifurcation for higher-order cost functions," *Appl. Phys. Exp.*, vol. 16, no. 1, Jan. 2023, Art. no. 014501, doi: [10.35848/1882-0786/acaba9](https://doi.org/10.35848/1882-0786/acaba9).
- [37] M. J. A. Schuetz, J. K. Brubaker, and H. G. Katzgraber, "Combinatorial optimization with physics-inspired graph neural networks," *Nature Mach. Intell.*, vol. 4, no. 4, pp. 367–377, Apr. 2022, doi: [10.1038/s42256-022-00468-6](https://doi.org/10.1038/s42256-022-00468-6).
- [38] C. Bick, E. Gross, H. A. Harrington, and M. T. Schaub, "What are higher-order networks?" *SIAM Rev.*, vol. 65, no. 3, pp. 686–731, Aug. 2023, doi: [10.1137/21M1414024](https://doi.org/10.1137/21M1414024).

- [39] Y. Shang, "Non-linear consensus dynamics on temporal hypergraphs with random noisy higher-order interactions," *J. Complex Netw.*, vol. 11, no. 2, pp. 509–512, Apr. 2023, doi: [10.1093/comnet/cnad009](https://doi.org/10.1093/comnet/cnad009).
- [40] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, "Networks beyond pairwise interactions: Structure and dynamics," *Phys. Rep.*, vol. 874, pp. 1–92, Aug. 2020, doi: [10.1016/j.physrep.2020.05.004](https://doi.org/10.1016/j.physrep.2020.05.004).
- [41] S. Saha, J. Gao, and R. Gerlach, "A survey of the application of graph-based approaches in stock market analysis and prediction," *Int. J. Data Sci. Anal.*, vol. 14, no. 1, pp. 1–15, Jun. 2022, doi: [10.1007/s41060-021-00306-9](https://doi.org/10.1007/s41060-021-00306-9).
- [42] Y. Yan, B. Wu, T. Tian, and H. Zhang, "Development of stock networks using part mutual information and Australian stock market data," *Entropy*, vol. 22, no. 7, p. 773, Jul. 2020, doi: [10.3390/e22070773](https://doi.org/10.3390/e22070773).
- [43] A. Bielinskyi, V. Soloviev, S. Hushko, A. Kiv, and A. Matviychuk, "High-order networks and stock market crashes," in *Proc. 10th Int. Conf. Monitor, Model. Manage. Emergent Economy*, 2022, pp. 134–144, doi: [10.5220/0011931900003432](https://doi.org/10.5220/0011931900003432).
- [44] S. Yarkoni, A. Plaat, and T. Back, "First results solving arbitrarily structured maximum independent set problems using quantum annealing," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2018, pp. 1–6, doi: [10.1109/CEC.2018.8477865](https://doi.org/10.1109/CEC.2018.8477865).
- [45] H. Goto, "Bifurcation-based adiabatic quantum computation with a nonlinear oscillator network," *Sci. Rep.*, vol. 6, no. 1, p. 21686, Feb. 2016, doi: [10.1038/srep21686](https://doi.org/10.1038/srep21686).
- [46] K. Tatsumura, R. Hidaka, M. Yamasaki, Y. Sakai, and H. Goto, "A currency arbitrage machine based on the simulated bifurcation algorithm for ultrafast detection of optimal opportunity," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Oct. 2020, pp. 1–5, doi: [10.1109/ISCAS45731.2020.9181114](https://doi.org/10.1109/ISCAS45731.2020.9181114).
- [47] K. Tatsumura, R. Hidaka, J. Nakayama, T. Kashimata, and M. Yamasaki, "Pairs-trading system using quantum-inspired combinatorial optimization accelerator for optimal path search in market graphs," *IEEE Access*, vol. 11, pp. 104406–104416, 2023, doi: [10.1109/ACCESS.2023.3316727](https://doi.org/10.1109/ACCESS.2023.3316727).
- [48] K. Tatsumura, R. Hidaka, J. Nakayama, T. Kashimata, and M. Yamasaki, "Real-time trading system based on selections of potentially profitable, uncorrelated, and balanced stocks by NP-hard combinatorial optimization," *IEEE Access*, vol. 11, pp. 120023–120033, 2023, doi: [10.1109/ACCESS.2023.3326816](https://doi.org/10.1109/ACCESS.2023.3326816).
- [49] A. Hagberg, P. Swart, and D. S. Chult, "Exploring network structure, dynamics, and function using networkx," in *Proc. 7th Python Sci. Conf.*, 2008, pp. 11–15. [Online]. Available: <https://www.osti.gov/biblio/960616>
- [50] L. Perron and V. Furnon. *OR-Tools (Version: 9.4.1874)*. Google. Accessed: Jun. 1, 2023 [Online]. Available: <https://developers.google.com/optimization/>
- [51] P. Miskolczi, "Note on simple and logarithmic return," *Appl. Stud. Agribusiness Commerce*, vol. 11, nos. 1–2, pp. 127–136, 2017, doi: [10.22004/ag.econ.265595](https://doi.org/10.22004/ag.econ.265595).
- [52] R. Boppana and M. M. Halldórsson, "Approximating maximum independent sets by excluding subgraphs," *BIT*, vol. 32, no. 2, pp. 180–196, Jun. 1992, doi: [10.1007/BF01994876](https://doi.org/10.1007/BF01994876).
- [53] M. Padberg and G. Rinaldi, "A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems," *SIAM Rev.*, vol. 33, no. 1, pp. 60–100, Mar. 1991, doi: [10.1137/1033004](https://doi.org/10.1137/1033004).
- [54] S. B. Seidman, "Clique-like structures in directed networks," *J. Social Biol. Struct.*, vol. 3, no. 1, pp. 43–54, 1980, doi: [10.1016/0140-1750\(80\)90019-6](https://doi.org/10.1016/0140-1750(80)90019-6).
- [55] Y. Shang, "Large cliques in a directed inhomogeneous random graph," *Int. J. Comput. Math.*, vol. 90, no. 3, pp. 445–456, Mar. 2013, doi: [10.1080/00207160.2012.735663](https://doi.org/10.1080/00207160.2012.735663).
- [56] W. F. Sharpe, "Mutual fund performance," *J. Bus.*, vol. 39, no. 1, pp. 119–138, 1966, doi: [10.1086/294846](https://doi.org/10.1086/294846).
- [57] D. K. Backus, A. W. Gregory, and C. I. Telmer, "Accounting for forward rates in markets for foreign currency," *J. Finance*, vol. 48, no. 5, pp. 1887–1908, Dec. 1993, doi: [10.1111/j.1540-6261.1993.tb05132.x](https://doi.org/10.1111/j.1540-6261.1993.tb05132.x).
- [58] J. Menchero, A. Morozov, and P. Shepard, "Global equity risk modeling," in *Handbook of Portfolio Construction*. New York, NY, USA: Springer, 2010, pp. 439–480, doi: [10.1007/978-0-387-77439-8_15](https://doi.org/10.1007/978-0-387-77439-8_15).



RYO HIDAKA received the B.E. and M.E. degrees in systems design and informatics from the Kyushu Institute of Technology, Japan, in 2006 and 2008, respectively. He joined Toshiba Corporation, in 2008. He was engaged in the development of main processors (2D-to-3D conversion and local dimming) for digital televisions, an image recognition processor called Visconti™, and host controllers for flash-memory cards. His current research interests include domain-specific computing, high-level synthesis design methodology, and the proof-of-concept study with FPGA devices.



YOHEI HAMAKAWA received the B.E. and M.E. degrees in advanced electronics and optical science from Osaka University, Japan, in 2000 and 2002, respectively. He joined Toshiba Corporation, in 2002. He was engaged in the development of image processing engines for digital televisions and TV products, and distributed computing algorithm in deep learning. His research interests include domain-specific computing, quantum computation, optimization in quantum circuit design, and their applications.



JUN NAKAYAMA received the Bachelor of Arts degree in economic and social studies from The University of Manchester, U.K., in 2008, and the Master of Business Administration (M.B.A.) degree in finance from Hitotsubashi University, Japan, in 2017. He was the Portfolio Manager of Nomura Asset Management Company Ltd., from 2008 to 2020. He was engaged in the development and management of quant-based funds. He joined the Corporate Research and Development Center, Toshiba Corporation, in 2020. He is also currently pursuing the Ph.D. degree with the Financial Strategy Program, Hitotsubashi University Business School. His research interests include quantitative investment strategies, quantum-inspired computing technology, and trading strategies with advanced technologies.



KOSUKE TATSUMURA (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electronics, information and communications engineering from Waseda University, Japan, in 2000, 2001, and 2004, respectively. After working as a Postdoctoral Fellow with Waseda University, he joined the Corporate Research and Development Center, Toshiba Corporation, in 2006. He is currently a Chief Research Scientist, leading a research team and several projects toward realizing innovative industrial systems based on cutting-edge computing technology. He was a member of the Emerging Research Devices (ERD) Committee in the International Technology Roadmap for Semiconductors (ITRS), from 2013 to 2015. He has been a Lecturer with Waseda University, since 2013. He was a Visiting Researcher with the University of Toronto, from 2015 to 2016. His research interests include domain-specific computing, quantum/quantum-inspired computing, and their applications. He received the Best Paper Award at IEEE International Conference on Field-Programmable Technology (FTP), in 2016.

• • •