BOUNDS FOR RANDOM BINARY QUADRATIC PROGRAMS*

KARTHIK NATARAJAN†, DONGJIAN SHI‡, AND KIM-CHUAN TOH§

Abstract. In this paper, we consider a binary quadratic program (BQP) with random objective coefficients. Given only information on the marginal distributions of the objective coefficients, we propose a tight bound on the expected optimal value of the random BQP. We show that the complexity of computing this bound does not increase substantially with respect to the complexity of solving the corresponding deterministic BQP. For the quadratic unconstrained binary optimization (QUBO) problem with nonnegative off-diagonal random entries, the bound is shown to be computable in polynomial time. We generalize the asymptotic bound for the random quadratic assignment problem from independent random variables to dependent random variables and propose a new closed-form bound on the expected optimal value of the quadratic k-cluster problem. We also provide polynomial time computable upper bounds on the expected optimal value for the NP-hard instances using the linear and semidefinite programming relaxation of the deterministic BQP. The semidefinite programming bound of the expected optimal value of the random max-cut problem inherits the approximation ratio of 0.878 from the semidefinite relaxation of the deterministic max-cut problem. Computational experiments on random QUBO problems and random quadratic knapsack problems provide evidence of the quality of the bounds. Overall, our results indicate that the new bound on the expected optimal value of the random BQP is attractive since it exploits many of the nice results and formulations of the deterministic BQP and is valid under limited distributional information.

Key words. binary quadratic program, bounds, semidefinite program

AMS subject classifications. 90C09, 90C22

1. Introduction.

For a given $n \times n$ real symmetric matrix $Q \in \mathcal{S}^n$, consider the binary quadratic program (BQP):

$$\beta(Q) \quad = \quad \max\left\{ x^T Q x \;\middle|\; x \in \mathcal{X} \subseteq \{0,1\}^n \right\}. \tag{1}$$

BQP of the form (1) has applications in diverse areas including computer-aided design (Jünger et al. [24]), solid-state physics (Barahona [3]), machine scheduling (Alidaee, Kochenberger and Ahmadian [2]), capital budgeting (Laughhunn [28]) and computer vision (Rother et. al. [35]). Graph problems such as the max-cut and the maximum clique problems can be reformulated as BQPs. The problem is hence known to be NP-hard. In special instances, this problem has shown to be solvable in polynomial time (see Picard and Ratliff [32]). A variety of heuristics (see Beasley [5], Glover, Kochenberger and Alidaee [18]), approximations based on second order conic programming and semidefinite programming (see Kim and Kojima [25], Ghaddar, Vera and Anjos [17], Lasserre [27]) and exact approaches using branch and bounds techniques (Carter [12], Pardalos and Rodgers [31], Helmberg and Rendl [23]) have also been proposed to solve BQPs.

*

Funding: The research of the first and second authors was partly funded by the SUTD-MIT International Design Center grant number IDG31300105 on 'Optimization for Complex Discrete Choice' and the MOE Tier 2 grant number MOE2013-T2-2-168 on 'Distributional Robust Optimization for Consumer Choice in Transportation Systems'.

†Engineering Systems and Design, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372. (karthik_natarajan@sutd.edu.sg).

‡Engineering Systems and Design, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372. (dongjian_shi@163.com).

§Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road, Singapore 119076. (mattohkc@nus.edu.sg).

While the deterministic BQP has been extensively studied, to the best of our knowledge the random version of this problem has received lesser attention with the quadratic assignment problem (see Burkard and Fincke [11]) and the quadratic knapsack problem (see Cheng, Delage and Lisser [13]) being notable exceptions. In this paper, we study the random binary quadratic program with a symmetric random matrix $\tilde{Q}$ and develop a bound on the expected optimal value. For a continuous distribution θ with support in Ω for the random matrix $\tilde{Q}$, the expected optimal objective value for the BQP (1) averaged over the possible realizations of the coefficients is given as:

$$E_\theta\left[\beta(\tilde{Q})\right] = \int_\Omega \left[\max_{x\in\mathcal{X}\subseteq\{0,1\}^n} x^T Q x\right] d\theta(Q). \tag{2}$$

For a discrete distribution, where $\tilde{Q}$ takes values in the set $\Omega = \{Q^{(1)}, \ldots, Q^{(m)}\}$, each occurring with a probability θ_j for $j \in [m] = \{1, \ldots, m\}$ with $\theta_j \geq 0$ and $\sum_j \theta_j = 1$, the expected optimal value is given as:

$$E_\theta\left[\beta(\tilde{Q})\right] = \sum_{j\in[m]} \left[\max_{x\in\mathcal{X}\subseteq\{0,1\}^n} x^T Q^{(j)} x\right] \theta_j. \tag{3}$$

Evaluating the expected values in both (2) and (3) are clearly challenging since we need to solve a set of NP-hard BQPs where the set might be of finite or infinite cardinality.

1.1. Literature Review. There is a significant stream of research in probabilistic combinatorial optimization that deals with evaluating the expected optimal value of random combinatorial optimization problems with a linear objective function. It is known from this literature that evaluating the expected optimal value can be hard even when the deterministic optimization problem is solvable in polynomial time. An example is the complexity of computing the expected longest path on a directed acyclic graph which is #P-complete for arc lengths that are independent, discrete random variables (see Hagstrom [22]), though the deterministic version of this problem is solvable in polynomial time. The hardness of the random version of the problem arises from the possibility that even with two point distributions, under the assumption of independence, the joint distribution has an exponential number of support points given by $m = 2^n$ where n is the number of arcs in the graph. This in turn might lead to an exponential number of optimal paths, one for each support point. Several explicit formulas have been developed for the expected optimal value of classical combinatorial optimization problems such as the linear assignment problem (see Aldous [1]), the minimum spanning tree problem (see Frieze [15]) and the traveling salesperson problem (see Beardwood et al. [4]) under simplifying assumptions on the distributions. Most of these formulas are asymptotic in nature and derived under the assumption that the objective coefficients are independent and identically distributed with either a uniform or an exponential distribution. Fewer results are known when the assumptions of independence and identical distributions are violated.

One such probability model that drops the assumption of independence where results have been found for the expected optimal value is the Fréchet class of distributions with fixed marginal distributions. In this class of distributions, either continuous or discrete marginal distributions of the objective coefficients are given but the dependency structure is not specified. The joint distribution might thus be independent, positively correlated or negatively correlated with the fixed marginals. This class of distributions has been extensively studied in the risk management literature to estimate bounds on aggregate risks of portfolios from the risk of the individual assets (see McNeil, Frey and Embrechts [29]). Bertsimas, Natarajan and Teo [6] building on the results of Meilijson and Nadas [30] showed that for the Fréchet class of distributions, it is possible to derive interesting bounds for several classical combinatorial optimization problems with a linear objective function. Specifically they showed that the bound is computable in polynomial time if the deterministic combinatorial optimization problem is solvable in polynomial time. This is in contrast to the results for independent distributions. Furthermore, they calculated the exact asymptotic bounds for combinatorial optimization problems such as the linear assignment, spanning tree, and traveling salesman problems, under knowledge of identical marginal distributions but without the assumption of independence. In this paper, we focus on the Fréchet class of distributions and derive a new bound for the expected optimal value of random binary quadratic programs.

2. Main Problem.

Consider a Fréchet class of probability distributions Θ for the random symmetric $\tilde{Q}$ with information on the marginal distributions only. Suppose the marginal distribution for the random term $\tilde{Q}_{ij}$ is θ_{ij} with support in Ω_{ij} for each $i, j \in [n]$. The joint distribution then lies in the Fréchet class of distributions denoted as $\theta \in \Theta(\theta_{ij}, i, j \in [n])$ with the given marginal distributions. Note that under our assumption of symmetric matrices, $\theta_{ji} = \theta_{ij}$ and in fact $Q_{ij} = Q_{ji}$ for each realization of the random matrix. Since the probability distribution is incompletely specified, we focus on the tightest bound that is generated by the multivariate distribution of the objective coefficients that maximizes the expected optimal objective value of the BQP over all distributions in the class Θ. Our problem of interest is thus defined as:

$$\phi \quad = \quad \sup_{\theta \in \Theta} E_\theta \left[\max_{x \in \mathcal{X} \subseteq \{0,1\}^n} x^T \tilde{Q} x \right]. \tag{4}$$

2.1. Motivation.

To motivate the study of such an upper bound, we consider the concept of the expected value of perfect information (EVPI) from stochastic programming. Consider a decision maker who needs to solve multiple instances of BQPs where each instance is obtained by a random perturbation of the objective coefficient matrix. A simple heuristic approach that the decision maker might adopt is to solve the BQP using the expected value of the random matrix $\tilde{Q}$ as follows:

$$\beta(E_\theta[\tilde{Q}]) \quad = \quad \max_{x \in \mathcal{X} \subseteq \{0,1\}^n} x^T \big(E_\theta[\tilde{Q}]\big) x. \tag{5}$$

However this solution will be sub-optimal in general. EVPI is defined as the price that the decision-maker would be willing to pay in order to gain access to perfect information. Thus EVPI measures the loss of optimality for the decision-maker from not knowing the exact objective coefficient realizations and is defined as:

$$\text{EVPI} \quad = \quad E_\theta \left[\beta(\tilde{Q}) \right] - \beta(E_\theta[\tilde{Q}]), \tag{6}$$

where the random instances are generated from the distribution θ. Calculating the EVPI has been of significant interest in the stochastic programming community. The following excerpt is taken from Section 4.5, Chapter 4 in the book by Birge and Louveaux [8], *"There has always been a strong interest in trying to have a better understanding of when the EVPI takes large values and when they take low values.*

A definite answer to this question would greatly simplify the practice of stochastic programming. Only those programs with large EVPI would require the solution of a stochastic program." However calculating the EVPI is hard in general. An upper bound on the EVPI in this case provides an estimate on the maximum price that the decision-maker would be willing to pay in order to gain access to perfect information. Such an upper bound on the EVPI is obtained by using the bound ϕ as follows:

$$\text{EVPI} \quad \leq \quad \phi - \beta(E_\theta[\tilde{Q}]), \tag{7}$$

and is valid for all distributions θ in this set Θ. Furthermore, in special cases if ϕ can be computed efficiently, it implies that the bound on the EVPI is efficiently computable and tight in the worst-case.

This brings us to the main contributions in this paper which are as follows:

(a) In Theorem 2, we develop a computationally implementable reformulation to compute ϕ that exploits the structure of the set of distributions Θ. Our reformulation builds on the results of Meilijson and Nadas [30] and Bertsimas, Natarajan and Teo [6] for linear combinatorial optimization problems. The key insight that helps extend the earlier results is that the objective function is linear in the random matrix $\tilde{Q}$ and quadratic in the decision variables x. Using the reformulation, we show that the bound ϕ for quadratic unconstrained binary optimization (QUBO) problems where the off-diagonal entries of the matrix $\tilde{Q}$ are nonnegative random variables is computable in polynomial time (see Theorem 5 and Corollary 6). Thus the complexity of computing ϕ does not increase substantially with respect to the complexity of solving the corresponding deterministic problem. This is unlike computing the expected optimal value under independent distributions which is often hard even if the corresponding deterministic optimization problem is easy.

(b) We develop new closed form bounds for the random quadratic assignment problem and quadratic k-cluster problems. A classical result of Burkard and Fincke [11] shows that for the quadratic assignment problem on an $n \times n$ bipartite graph when the objective coefficients are random variables, independently distributed in $[0, 1]$, with probability tending to one, the ratio between the best and worst objective function values approaches one as the size of the problem tends to infinity. Extensions of this result to more general distributions have been developed in Szpankowski [39]. In Theorem 7, we derive new bounds for the quadratic assignment problem by dropping the assumption of independence and using only identical marginal distributional information. The bounds imply that there exists dependent distributions for which the random instances of QAP might be drastically different from the random instances generated from an independent distribution. We develop closed form bounds on the expected value of the quadratic assignment and the quadratic k-cluster problem for uniform, exponential (light-tailed) and Pareto (heavy-tailed) random variables. To the best of our knowledge, no closed form bounds are currently known for the quadratic k-cluster even under independent distributions. Our results make extensive use of symmetry in the feasible region of the quadratic assignment problem and quadratic k-cluster problems.

(c) Building on the linear and semidefinite programming relaxations for the deterministic BQP, we develop polynomial time computable upper bounds on ϕ. The corresponding semidefinite programming bound on the expected optimal value of the random max-cut inherits the approximation ratio of 0.878 from the semidefinite relaxation of the deterministic max-cut problem. We also provide a subgradient method such that each step involves solving a BQP to numerically approximate the exact bound ϕ. We provide numerical results on random QUBO and quadratic knapsack problems that validate the quality of the bounds.

3. Reformulation for computing the bound ϕ.

In this section, we propose a reformulation to compute the tight upper bound on the expected optimal value of a random BQP (4) that exploits the assumption that only the marginal distributions of $\tilde{Q}$ are given. Tightness of the upper bound refers to the existence of a joint distribution with the given marginal distributions that attains the bound or a sequence of joint distributions that attains the upper bound in the limit. Our approach is based on the results in Meilijson and Nadas [30] who developed a convex majorization approach to compute the upper bound on the expected length of a critical path in a project network for the Fréchet class of distributions. Weiss [42] generalized this bound to linear combinatorial optimization problems such as the shortest path, maximum flow, and the reliability problem. Extensions of this approach to limited information on the marginal distributions have been proposed in Birge and Maddox [9] and Bertsimas, Natarajan and Teo [6, 7] among others. The main result in these papers is outlined in the next lemma.

Lemma 1. *(Meilijson and Nadas [30]) Assume that the marginal distributions of $\tilde{c}$ are given as θ_i, $i \in [n]$ with finite second moments and support in Ω_i, $i \in [n]$. The joint distribution θ of $\tilde{c}$ lies in a set Θ, where $\Theta = \Theta(\theta_i, i \in [n])$ represents the class of distributions with the prescribed marginal distributions for $\tilde{c}$. Then the tight upper bound on the expected optimal value of the random combinatorial optimization problem with a linear objective function:*

$$\sup_{\theta \in \Theta} E_\theta \left[\max_{x \in \mathcal{X} \subseteq \{0,1\}^n} \tilde{c}^T x \right], \tag{8}$$

is given by the optimal objective value to the following problem:

$$\inf_{d \in \Re^n} \left[\max_{x \in \mathcal{X} \subseteq \{0,1\}^n} d^T x + \sum_{i \in [n]} \mathbb{E}_{\theta_i} (\tilde{c}_i - d_i)^+ \right], \tag{9}$$

where $y^+ = \max(0, y)$.

The reformulation in (9) can be interpreted as finding a deterministic estimate d of the random objective vector $\tilde{c}$ such that the objective function of the combinatorial optimization problem given by $\max_{x \in \mathcal{X}} d^T x$ is balanced with a penalty term given by $\sum_i \mathbb{E}_{\theta_i}(\tilde{c}_i - d_i)^+$ for estimating d differently from the random $\tilde{c}$. This result can be extended to bounds on the expected optimal value of a random BQP for the Fréchet class of distributions as we discuss next by observing that the objective function is linear in the entries of the matrix $\tilde{Q}$.

Theorem 2. *Assume that the marginal distributions of the symmetric matrix $\tilde{Q}$ are given as $\theta_{ij}, i, j \in [n]$ with finite second moments and support in $\Omega_{ij}, i, j \in [n]$ with $\theta_{ij} = \theta_{ji}$ and $\Omega_{ij} = \Omega_{ji}$. The joint distribution θ of $\tilde{Q}$ lies in the set Θ, where $\Theta = \Theta(\theta_{ij}, i, j \in [n])$ represents the class of distributions with the prescribed marginal distributions for $\tilde{Q}$. Then the tight upper bound on the expected optimal value of the* *random BQP in (4) is given by the optimal objective value to the following problem:*

$$\phi = \inf_{R \in \mathcal{S}^n} \left[\max_{x \in \mathcal{X} \subseteq \{0,1\}^n} x^T R x + \sum_{i \in [n]} \sum_{j \in [n]} E_{\theta_{ij}} (\tilde{Q}_{ij} - R_{ij})^+ \right]. \tag{10}$$

Proof. Notice that the BQP in (1) is equivalent to the binary optimization problem with a linear objective function:

$$\max_{X \in \mathcal{Y}} \sum_{i \in [n]} \sum_{j \in [n]} Q_{ij} X_{ij},$$

where $\mathcal{Y} = \{X = xx^T \mid x \in \mathcal{X} \subseteq \{0,1\}^n\}$. Hence (4) can be written as

$$\phi = \sup_{\theta \in \Theta} E_\theta \left[\max_{X \in \mathcal{Y}} \sum_{i \in [n]} \sum_{j \in [n]} \tilde{Q}_{ij} X_{ij} \right].$$

Since $\mathcal{X} \subseteq \{0,1\}^n$, we have $\mathcal{Y} \subseteq \{0,1\}^{n \times n}$. By Lemma 1 with $X \in \mathcal{Y} \subseteq \mathcal{S}^n$, we obtain:

$$\phi = \inf_{R \in \mathcal{S}^n} \left[\max_{X \in \mathcal{Y}} \sum_{i \in [n]} \sum_{j \in [n]} R_{ij} X_{ij} + \sum_{i \in [n]} \sum_{j \in [n]} E_{\theta_{ij}} (\tilde{Q}_{ij} - R_{ij})^+ \right],$$

$$= \inf_{R \in \mathcal{S}^n} \left[\max_{x \in \mathcal{X}} x^T R x + \sum_{i \in [n]} \sum_{j \in [n]} E_{\theta_{ij}} (\tilde{Q}_{ij} - R_{ij})^+ \right].$$

In Theorem 2, we assumed that the full marginal distributional information of $\tilde{Q}$ is given. The result extends to the case when only partial information of the marginal distributions is available. Consider for example, that only the mean and variance for the marginal distributions of $\tilde{Q}$ are given.

THEOREM 3. *Assume that the marginal distributions for the symmetric matrix $\tilde{Q}$ lie in the sets Θ_{ij}, $i, j \in [n]$ with given mean and variance information:*

$$\Theta_{ij} = \left\{ \theta_{ij} \;\middle|\; E_{\theta_{ij}}(\tilde{Q}_{ij}) = \mu_{ij}, \quad E_{\theta_{ij}}(\tilde{Q}_{ij}^2) = \mu_{ij}^2 + \sigma_{ij}^2 \right\}, \; i, j \in [n].$$

The joint distribution θ of $\tilde{Q}$ lies in the set Θ, where $\Theta = \Theta(\theta_{ij}, i, j \in [n])$ represents the class of distributions with prescribed marginal moments for each $\tilde{Q}_{ij}$. Then the tight upper bound for the expected optimal value of the random BQP in (4) is given by the optimal objective value to the following problem:

$$\phi = \inf_{R \in \mathcal{S}^n} \left[\max_{x \in \mathcal{X}} x^T R x + \frac{1}{2} \sum_{i \in [n]} \sum_{j \in [n]} (\mu_{ij} - R_{ij}) + \sqrt{(\mu_{ij} - R_{ij})^2 + \sigma_{ij}^2} \right]. \tag{11}$$

Proof. This result can be shown through the following set of equalities and inequalities:

$$
\begin{aligned}
\phi &= \sup_{\theta_{ij}\in\Theta_{ij},i,j\in[n]} \ \sup_{\theta\in\Theta(\theta_{ij},i,j\in[n])} E_\theta\left[\max_{x\in\mathcal{X}} x^T\tilde{Q}x\right],\\
&= \sup_{\theta_{ij}\in\Theta_{ij},i,j\in[n]} \ \inf_{R\in\mathcal{S}^n}\left[\max_{x\in\mathcal{X}} x^T Rx + \sum_{i\in[n]}\sum_{j\in[n]} E_{\theta_{ij}}(\tilde{Q}_{ij}-R_{ij})^+\right],\\
&\le \inf_{R\in\mathcal{S}^n}\left[\max_{x\in\mathcal{X}} x^T Rx + \sum_{i\in[n]}\sum_{j\in[n]} \sup_{\theta_{ij}\in\Theta_{ij}} E_{\theta_{ij}}(\tilde{Q}_{ij}-R_{ij})^+\right],\\
&= \inf_{R\in\mathcal{S}^n}\left[\max_{X\in\mathcal{Y}} \sum_{i\in[n]}\sum_{j\in[n]} R_{ij}X_{ij} + \sum_{i\in[n]}\sum_{j\in[n]} \sup_{\theta_{ij}\in\Theta_{ij}} E_{\theta_{ij}}(\tilde{Q}_{ij}-R_{ij})^+\right],\\
&= \sup_{\theta\in\Theta(\Theta_{ij},i,j\in[n])} E_\theta\left[\max_{X\in\mathcal{Y}} \sum_{i\in[n]}\sum_{j\in[n]} \tilde{Q}_{ij}X_{ij}\right],\\
&= \phi,
\end{aligned}
$$

where the second equality comes from Theorem 2, the third inequality comes from the interchange of the supremum and the infimum, the fourth equality comes from the linearization of the objective while the fifth equality comes from Theorem 3.1 in Bertsimas, Natarajan and Teo [6] who showed tightness of the bound for linear objective functions with two moment information. When only the mean and variance of $\tilde{Q}_{ij}$ are given, $\sup_{\theta_{ij}\in\Theta_{ij}} E_{\theta_{ij}}(\tilde{Q}_{ij}-R_{ij})^+$ has a simple closed-form expression based on the Cauchy-Schwarz inequality (see Scarf [36]):

$$
\sup_{\theta_{ij}\in\Theta_{ij}} E_{\theta_{ij}}(\tilde{Q}_{ij}-R_{ij})^+ = \frac{1}{2}\left[(\mu_{ij}-R_{ij}) + \sqrt{(\mu_{ij}-R_{ij})^2+\sigma_{ij}^2}\right],
$$

resulting in the formulation (11). □

3.1. Polynomially solvable instances of ϕ. In this section, we discuss an important implication of the formulations in Theorems 2 and 3 by identifying cases when the bound ϕ is computable in polynomial time. The formulations exploit the marginal distributional information structure to provide a finite dimensional optimization reformulation over a symmetric matrix variable R where the objective function consists of two parts:

(a) Optimal objective value of a deterministic BQP with R as the coefficient matrix,
(b) Sum of univariate expectations of the functions of the form $(\tilde{Q}_{ij}-R_{ij})^+$.

It is easy to see that for deterministic problems, the bound ϕ simply reduces to $\beta(Q)$.

LEMMA 4. *Assume that the matrix $\tilde{Q} = Q$ with probability 1. Then the tight upper bound on the expected optimal value of the BQP in (4) reduces to the optimal value $\beta(Q)$ of the deterministic BQP.*

Proof. For a deterministic matrix Q, the bound ϕ in (10) reduces to:

$$
\phi = \inf_{R\in\mathcal{S}^n}\left[\max_{x\in\mathcal{X}\subseteq\{0,1\}^n} x^T Rx + \sum_{i\in[n]}\sum_{j\in[n]}(Q_{ij}-R_{ij})^+\right]. \tag{12}
$$

It is easy to verify that in this case, the optimal $R = Q$. If not and say $R_{ij} < Q_{ij}$, then by increasing R_{ij} to Q_{ij}, we decrease the second term in (12) at a rate of 1 while increasing the first term at most at a rate of 1. Similarly if $R_{ij} > Q_{ij}$, then by decreasing R_{ij} to Q_{ij}, the second term remains unchanged while the first term does not increase and possibly decreases at a rate of 1. ∎

Note that clearly the bounds in Theorems 2 and 3 are NP-hard to compute, since the deterministic BQP is already NP-hard. However the objective function has a useful separable structure that helps preserve polynomial time computability of the bound ϕ when the deterministic BQP is solvable in polynomial time.

THEOREM 5. *Assume that the following two conditions hold:*

(a) The deterministic BQP is solvable in polynomial time,

(b) For any $i, j \in [n]$ and a given value of R_{ij}, the univariate probability $P_{\theta_{ij}}(\tilde{Q}_{ij} \leq R_{ij})$ and the expectation $E_{\theta_{ij}}(\tilde{Q}_{ij} - R_{ij})^+$ is efficiently computable.

Then, the bound ϕ is computable in polynomial time.

Proof. By introducing new variables t and S_{ij} for $i, j \in [n]$, we can reformulate (10) as:

$$\phi = \min \left\{ t + \sum_{i\in[n]} \sum_{j\in[n]} S_{ij} \;\middle|\; t \geq x^T R x, \forall x \in \mathcal{X}, S_{ij} \geq E_{\theta_{ij}}(\tilde{Q}_{ij} - R_{ij})^+, \forall i, j \in [n] \right\}.$$

Consider the separation version of this problem: Given t^*, S^* and R^*, check if it is feasible, if not find a hyperplane which separates the point from the feasible region efficiently. We next argue that under assumptions (a) and (b), the separation problem is efficiently solvable. Under assumption (a), we can check for the feasibility of the first constraint in polynomial time by solving the deterministic BQP with objective coefficient matrix R^*. Let $x^* \in \mathcal{X}$ be an optimal solution for the problem. If $t^* \geq x^{*T} R^* x^*$, the constraint is feasible else we find a separating hyperplane with the optimal solution x^*. Similarly, we can check for the feasibility of the second set of constraints in polynomial time by evaluating the n^2 expected values of convex functions. For each of these cases, we verify if $S^*_{ij} \geq E_{\theta_{ij}}(\tilde{Q}_{ij} - R^*_{ij})^+$ and if not identify the separating hyperplane by computing a subgradient for the convex function $E_{\theta_{ij}}(\tilde{Q}_{ij} - R_{ij})^+$ at R^*_{ij}. Under the assumption that the event $\tilde{Q}_{ij} = R_{ij}$ has zero probability, the gradient is given by $-P_{\theta_{ij}}(\tilde{Q}_{ij} > R_{ij})$. More generally if the event has nonzero probability, the subdifferential is given by the interval $[-P_{\theta_{ij}}(\tilde{Q}_{ij} \geq R_{ij}), -P_{\theta_{ij}}(\tilde{Q}_{ij} > R_{ij})]$ (see Chapter 7 in Shapiro, Dentcheva and Ruszczyński [37]). From the equivalence of separation and optimization, the problem is solvable in polynomial time under assumptions (a) and (b). ∎

This result generalizes the earlier results of Meilijson and Nadas [30] and Bertsimas, Natarajan and Teo [6] from a linear objective function to the quadratic objective function. As an implication of this result, consider the quadratic unconstrained binary optimization (QUBO) problem:

$$\beta_{\text{qubo}}(Q) \quad = \quad \max \left\{ x^T Q x \mid x \in \{0, 1\}^n \right\}, \tag{13}$$

and the corresponding bound on the expected optimal value:

$$\phi_{\text{qubo}} \quad = \quad \sup_{\theta\in\Theta} E_\theta \left[\beta_{\text{qubo}}(\tilde{Q}) \right]. \tag{14}$$

A classical result by Picard and Ratliff [32] shows that under the assumption that the off-diagonal entries of the matrix Q are nonnegative, the deterministic QUBO problem is solvable in polynomial time. Applying Theorem 5, we then obtain the following result.

COROLLARY 6. *Assume that the off-diagonal entries of the matrix $\tilde{Q}$ are nonnegative random variables where the marginal distributions of these entries satisfy assumption (b) in Theorem 5. This includes the uniform distribution in $[0,1]$, the exponential distribution and discrete distributions with nonnegative support as special cases. In this case, the bound ϕ_{qubo} in (14) for the Fréchet class of distributions is computable in polynomial time.*

4. Closed form instances of ϕ. In this section, we derive closed-form expressions for the bound for two instances of BQPs under the assumption that the objective coefficients are random and identically distributed, but not necessarily independent.

4.1. Quadratic assignment problem. We first consider the quadratic assignment problem (QAP) where given coefficients Q_{ijkl} for $i,j,k,l \in [n]$, the goal is to determine an assignment that maximizes the quadratic objective function:

$$\beta_{\text{qap}}(Q) \quad = \quad \max \quad \left\{ \sum_{i,j,k,l \in [n]} Q_{ijkl} x_{ij} x_{kl} \;\middle|\; x \in \mathcal{X}_{\text{qap}} \right\}. \tag{15}$$

where:

$$\mathcal{X}_{\text{qap}} \quad = \quad \left\{ x_{ij} \in \{0,1\}, \forall i,j \in [n] \;\middle|\; \sum_j x_{ij} = 1, \forall i \in [n], \sum_i x_{ij} = 1, \forall j \in [n] \right\}.$$

The tight upper bound on the expected optimal value of the random QAP is given as:

$$\phi_{\text{qap}} \quad = \quad \sup_{\theta \in \Theta} E_\theta \left[\beta_{\text{qap}}(\tilde{Q}) \right]. \tag{16}$$

The feasible region of the QAP in (4.1) denoted by $\mathcal{X}_{\text{qap}}$ has the following characteristics: the number of variables is n^2, the number of feasible solutions is $n!$, each feasible assignment has n variables set to 1, each variable x_{ij} takes a value of 1 in $(n-1)!$ solutions while each pair of variables x_{ij} and x_{kl} for $i \neq k$ and $j \neq l$ takes a value of 1 jointly in $(n-2)!$ solutions. We derive a new bound on the expected value of the QAP problem for the Fréchet class of distributions with identical distributions in the next theorem.

THEOREM 7. *Let $\tilde{Q}_{ijkl}$ for $i,j,k,l \in [n]$ be absolutely continuous random variables with an identical cumulative distribution function $\tilde{q} \sim F(\cdot)$ with finite expected value. For $n \geq 2$, the bound for the expected optimal value of the random QAP in (16) is given by:*

$$\phi_{qap} = nE\left(\tilde{q} \;\middle|\tilde{q} \geq F^{-1}\left(1 - \frac{1}{n}\right)\right) + n(n-1)E\left(\tilde{q} \;\middle|\tilde{q} \geq F^{-1}\left(1 - \frac{1}{n(n-1)}\right)\right). \tag{17}$$

Proof. The Karush-Kuhn-Tucker optimality conditions for problem (10) for QAP is given as:

(i) $\lambda_x \geq 0, \forall x \in \mathcal{X}_{\text{qap}}$,

(ii) $\sum_{x \in \mathcal{X}_{\text{qap}}} \lambda_x = 1$,

(iii) $P_{\theta_{ijkl}}(\tilde{Q}_{ijkl} \geq R_{ijkl}) = \sum_{x \in \mathcal{X}_{\text{qap}}: x_{ij}=1, x_{kl}=1} \lambda_x, \forall i, j, k, l \in [n]$,

(iv) $\lambda_x \left(\max_{y \in \mathcal{X}_{\text{qap}}} y^T R y - x^T R x \right) = 0, \forall x \in \mathcal{X}_{\text{qap}}$.

We now identify a solution that satisfies the Karush-Kuhn-Tucker optimality conditions. Set $\lambda_x = 1/n!$ which satisfies conditions (i) and (ii) since the feasible region for QAP consists of $n!$ solutions. Condition (iii) reduces to the following three cases:

(i) For $i \neq k$, $j \neq l$, we have $P_{\theta_{ijkl}}(\tilde{Q}_{ijkl} \geq R_{ijkl}) = \frac{(n-2)!}{n!}$. Hence $R_{ijkl} = F^{-1}(1 - \frac{1}{n(n-1)})$.

(ii) For $i = k$, $j = l$, we have $P_{\theta_{ijkl}}(\tilde{Q}_{ijkl} \geq R_{ijkl}) = \frac{(n-1)!}{n!}$. Hence $R_{ijkl} = F^{-1}(1 - \frac{1}{n})$.

(iii) For all other values of i, j, k, l, we have $P_{\theta_{ijkl}}(\tilde{Q}_{ijkl} \geq R_{ijkl}) = 0$. Hence R_{ijkl} can be set to $F^{-1}(1)$.

Plugging in the values, for every feasible solution $x \in \mathcal{X}_{\text{qap}}$, we have $x^T R x = nF^{-1}(1 - \frac{1}{n}) + n(n-1)F^{-1}(1 - \frac{1}{n(n-1)})$. This satisfies the complementarity condition in (iv) since all the solutions attain the same objective value. Substituting the variables into the objective function, we get:

□

$$
\begin{aligned}
\phi_{\text{qap}} &= \max_{x \in \mathcal{X}_{\text{qap}}} x^T R x + \sum_{i \in [n]} \sum_{j \in [n]} \sum_{k \in [n]} \sum_{l \in [n]} E_{\theta_{ijkl}} (\tilde{Q}_{ijkl} - R_{ijkl})^+, \\
&= nF^{-1}\left(1 - \frac{1}{n}\right) + n(n-1)F^{-1}\left(1 - \frac{1}{n(n-1)}\right) \\
&\quad + n^2 E\left(\tilde{q} - F^{-1}\left(1 - \frac{1}{n}\right)\right)^+ \\
&\quad + n^2(n-1)^2 E\left(\tilde{q} - F^{-1}\left(1 - \frac{1}{n(n-1)}\right)\right)^+, \\
&= nE\left(\tilde{q} \;\middle|\; \tilde{q} \geq F^{-1}\left(1 - \frac{1}{n}\right)\right) \\
&\quad + n(n-1)E\left(\tilde{q} \;\middle|\; \tilde{q} \geq F^{-1}\left(1 - \frac{1}{n(n-1)}\right)\right).
\end{aligned}
$$

Burkard and Fincke [11] show that for the quadratic assignment problem on an $n \times n$ bipartite graph when the objective coefficients are independent random variables distributed in $[0, 1]$, with probability tending to one, the ratio between the best and worst objective function values approaches one as the size of the problem tends to infinity. This implies that asymptotically almost every algorithm finds the optimal solution for quadratic assignment problem instances generated under the independence assumption. This is surprising since the deterministic version of the problem is hard. Szpankowski [39] extended this result to more general independent and identically distributed random variables with finite first three moments. Several other authors including Frenk, Van Houweninge and Rinnooy Kan [14] and Rhee [34] have also studied the stochastic version of the quadratic assignment problem under independence and shown that the optimal value is of the form $n^2\mu + O(n^{3/2}\sqrt{\log n})$ where $\mu = E(\tilde{Q}_{ijkl})$. Using Theorem 7, we now provide new bounds for the quadratic assignment problem under arbitrary dependence among the random variables. In Table 1, we characterize the bounds for the Fréchet class of distributions with the following marginals - the uniform random in $[0, 1]$ which is a bounded random variable, the exponential distribution which is a light-tailed distribution and a Pareto distribution which is a heavy-tailed distribution.

Table 1: Upper bound on expected value of quadratic assignment problem

Distribution	$F(x)$	ϕ_{qap}
Uniform[0,1]	x for $x \in [0,1]$	$n^2 - 1$
Exponential(1)	$1 - e^{-x}$ for $x \geq 0$	$n^2 \log(n(n-1)) + n^2 - n\log(n-1)$
Pareto($\alpha > 1$)	$1 - x^{-\alpha}$ for $x \geq 1$	$\frac{\alpha}{\alpha - 1} n^{1+1/\alpha}\left((n-1)^{1+1/\alpha} + 1\right)$

Unlike the independence model for the uniform [0,1] random variable where the expected optimal value scales as $n^2/2$ (the average case), by dropping the assumption of independence, Table 1 shows that the maximum expected value is $n^2 - 1$ (the extreme case). This implies that the asymptotic behavior for the random QAP when the assumption of independence is not satisfied can be drastically different and the observation that almost any feasible solution for QAP is optimal can break down. Similarly, the bound scales as $\Theta(n^2 \log n)$ for the exponential distribution and $\Theta(n^{\frac{2(\alpha+1)}{\alpha}})$ for the Pareto distribution. As should be expected, for the heavy tailed Pareto distribution with α close to 1, the bounds for the optimal value scale more rapidly.

4.2. Quadratic k-cluster problem. The key insight that is used in deriving the closed-form bound for the QAP in Theorem 7 is to exploit the identical distribution assumption and the high degree of symmetry that the feasible region of QAP possesses. We can use this insight to also generate bounds for other binary quadratic programs such as the quadratic k-cluster problem as we show next.

Given a set of nodes $[n]$, the quadratic k-cluster (QKC) problem is to determine a subset $\mathcal{S} \subseteq [n]$ of k nodes such that the sum of the weights of the edges between nodes in $\mathcal{S}$ is maximized. Let Q_{ij} denote the weight of the edge between node i and j with $Q_{ij} = Q_{ij}$ and Q_{ii} denote the weight of the self-loop for node i. The quadratic k-cluster problem is formulated as the binary quadratic program:

$$\beta_{\text{qkc}}(Q) = \max \left\{ \sum_{i\in[n]} \sum_{j\in[n]} Q_{ij} x_i x_j \;\middle|\; \sum_{i\in[n]} x_i = k, x_i \in \{0,1\}, \forall i \in [n] \right\}, \tag{18}$$

where $x_i = 1$ if node i is selected in the cluster and 0 otherwise. The feasible region of the QKC in (18) has the following characteristics: the number of variables is n, the number of feasible solutions is $\binom{n}{k}$, each feasible solution has k variables set to 1, each variable x_i takes a value of 1 in $\binom{n-1}{k-1}$ solutions while each pair of variables x_i and x_j for $i \neq j$ takes a value of 1 jointly in $\binom{n-2}{k-2}$ solutions. Consider the problem of finding the tight upper bound on the expected optimal value of the random QKC problem:

$$\phi_{\text{qkc}} \quad = \quad \sup_{\theta\in\Theta} E_\theta \left[\beta_{\text{qkc}}(\tilde{Q})\right]. \tag{19}$$

Using an approach similar to the proof of Theorem 7 (we drop the details here), we obtain the following result.

COROLLARY 8. *Let $\tilde{Q}_{ij}$ for $i, j \in [n]$ be absolutely continuous random variables with an identical cumulative distribution function $\tilde{q} \sim F(\cdot)$ with finite expected value.*

The bound for the expected optimal value of the random QKC in (19) for $k \geq 2$ is given by:

$$\phi_{qkc} = kE\left(\tilde{q}\ \left|\tilde{q} \geq F^{-1}\left(1-\frac{k}{n}\right)\right.\right) + k(k-1)E\left(\tilde{q}\ \left|\tilde{q} \geq F^{-1}\left(1-\frac{k(k-1)}{n(n-1)}\right)\right.\right). \tag{20}$$

In Table 2, we characterize the bounds for the Fréchet class of distributions for the quadratic k-cluster problem with the uniform, exponential and Pareto random variables.

Table 2: Upper bound on expected value of quadratic k-cluster problem for $k \geq 2$

Distribution	ϕ_{qkc}
Uniform[0,1]	$k^2\left(1-\frac{1}{2n}-\frac{(k-1)^2}{2n(n-1)}\right)$
Exponential(1)	$k^2\log\left(\frac{n(n-1)}{k(k-1)}\right)+k^2-k\log\left(\frac{n-1}{k-1}\right)$
Pareto($\alpha > 1$)	$\frac{\alpha}{\alpha-1}n^{1/\alpha}k^{1-1/\alpha}\left(1+(n-1)^{1/\alpha}(k-1)^{1-1/\alpha}\right)$

We now use these bounds to compute an upper bound on EVPI as discussed in (7). For the quadratic k-cluster problem, this bound is computed as follows:

$$\text{EVPI}_{\text{qkc}} \quad \leq \quad \phi_{\text{qkc}} - \beta_{\text{qkc}}(E_\theta[\tilde{Q}]). \tag{21}$$

For example with uniform $[0,1]$ marginals, this bound reduces to:

$$\text{EVPI}_{\text{qkc}} \quad \leq \quad \frac{k^2}{2}\left(1-\frac{1}{n}-\frac{(k-1)^2}{n(n-1)}\right). \tag{22}$$

The bounds on the EVPI for the three distributions is plotted in Figure 1. From the figure, we can observe that the maximum value of the EVPI bound for $n = 25$ is at $k = 18$ for the uniform distribution, $k = 15$ for the exponential distribution and $k = 13$ for the Pareto distribution with $\alpha = 2$. In other words for $n = 25$, a decision-maker would benefit most from obtaining information on the samples of the worst-case joint distribution with uniform $[0,1]$ marginals for $k = 18$. Thus, the bounds provide useful information to help identify the value of k at which the EVPI is maximum in the worst-case for a given marginal distribution. We also simulated the EVPI by generating random samples from the i.i.d. distribution with the prescribed marginals. For each value of k, we solved 100 random instances of the quadratic k-cluster problem to estimate the EVPI. Simulating the EVPI is however much more computationally expensive in comparison to the closed form nature of the worst-case bounds. As the figures indicate, the EVPI is much higher for the heavy tailed Pareto distribution in comparison to the other two distributions, indicating the effect of the marginal distribution on the expected optimal value. Furthermore, the figure illustrates that there is a fairly significant gap between the upper bound on the EVPI which is valid across all possible distributions with the given marginals and the independent distribution. To better illustrate that this is because the worst-case distribution is highly correlated, we first provide a characterization of an extremal distribution $\theta^* \in \Theta$ that attains the upper bound for uniform marginals in a two step process as follows:

(a) Generate each solution $x \in \mathcal{X}_{\text{qkc}}$ where $\mathcal{X}_{\text{qkc}}$ is the feasible region in (18) with probability $\lambda = 1/\binom{n}{k}$.

(b) Given the solution $x \in \mathcal{X}_{\text{qkc}}$, generate the random matrix $\tilde{Q}$ as follows:

1. For each $i \in [n]$, if $x_i = 1$, generate $\tilde{Q}_{ii} \sim n\tilde{u}_{ii}\mathbf{1}(\tilde{u}_{ii} \geq 1 - k/n)/k$ else if $x_i = 0$, generate $\tilde{Q}_{ii} \sim n\tilde{u}_{ii}\mathbf{1}(\tilde{u}_{ii} \leq 1 - k/n)/(n-k)$ where $\tilde{u}_{ii} \sim \tilde{U}[0,1]$ are independent uniform random variables in $[0,1]$ and $\mathbf{1}(\cdot)$ is the indicator function.
2. For each $i \neq j$, if $x_i x_j = 1$, then generate $\tilde{Q}_{ij} \sim n(n-1)\tilde{u}_{ij}\mathbf{1}(\tilde{u}_{ij} \geq 1 - k(k-1)/n(n-1))/k(k-1)$ else if $x_i x_j = 0$, generate $\tilde{Q}_{ij} \sim n(n-1)\tilde{u}_{ij}\mathbf{1}(\tilde{u}_{ij} \leq 1 - k(k-1)/n(n-1))/(n(n-1) - k(k-1))$ where $\tilde{u}_{ij} \sim \tilde{U}[0,1]$ are independent uniform random variables in $[0,1]$.
3. Set $\tilde{Q} = (\tilde{Q} + \tilde{Q}^T)/2$.

A careful numerical analysis of the dependency structure in the extremal distribution θ^* indicates that the diagonal entries of the random matrix are negatively correlated with each other and they become more negatively correlated for moderate values of k as compared to extreme values of k. Furthermore, the entries $\tilde{Q}_{ij}$ and $\tilde{Q}_{ik}$ along a given row are positively correlated with the random variables becoming more positively correlated for moderate values of k as compared to extreme values of k. Inspired by the structure of this extremal distribution, we also generate a fixed dependent distribution in the set Θ as follows:

(a) Generate the diagonal entries of the $\tilde{Q}$ matrix with a Gaussian copula with correlation set to $-1/(n-1)$.

(b) Generate the off-diagonal entries $\tilde{Q}_{ij}$ of the ith row in the $\tilde{Q}$ matrix with a comonotonic copula (perfectly correlated) by setting it equal to Q_{ii}.

(c) Symmetrize the matrix with $\tilde{Q} = (\tilde{Q} + \tilde{Q}^T)/2$.

Note that unlike the extremal distribution where the correlation structure changes with k, in this dependent distribution, the correlation structure is fixed. In Figure 1, we plot the simulated EVPI for this dependent distribution with uniform marginals using 100 random instances for each k. The figure clearly illustrates that the EVPI upper bound is much closer to the simulated EVPI for this dependent distribution as compared to the independent distribution.

5. Upper bounds on ϕ.

In this section, we propose upper bounds $\bar{\phi}$ on ϕ that are polynomial time computable by building on the linear and semidefinite programming relaxations of the deterministic binary quadratic program. We show that the semidefinite programming upper bound for the random max-cut problem preserves the well-known Goemans-Williamson approximation guarantee of the deterministic max-cut problem. We also develop a subgradient method to compute a numerical approximation of the bound ϕ by solving a set of BQPs. Numerical results validate the quality of the bounds.

5.1. Linear and semidefinite programming bounds on ϕ.

Recollect that in formulation (10), the first term in the objective function is $\beta(R) = \max\{x^T Rx \mid x \in \mathcal{X}\}$ which is NP-hard to compute for a given R and hence computing the bound ϕ is hard in general unless P = NP. We now propose a polynomial time computable upper bound on ϕ by using the linear programming relaxation of the feasible region. Assume the feasible region for the BQP is given as:

$$\mathcal{X} = \{x \in \{0,1\}^n \mid Ax = b\}, \tag{23}$$

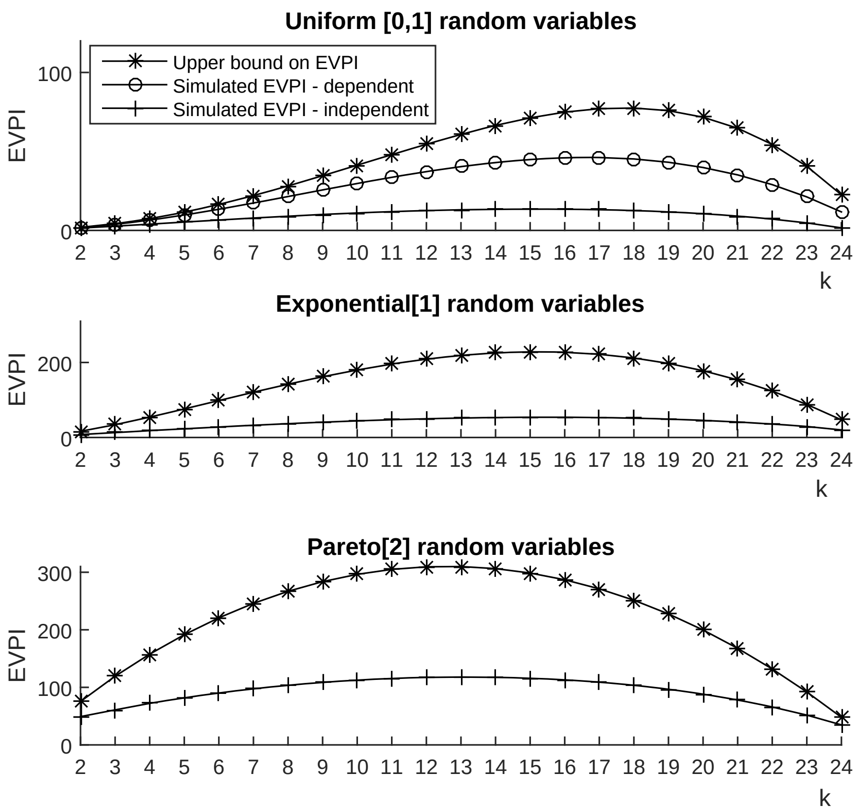

Fig. 1: EVPI for quadratic k-cluster problem for $n = 25$.

where $A \in \Re^{m\times n}$ and $b \in \Re^n$. Using the standard linearization approach to tackle the product of binary variables $x_i x_j$, we can reformulate the deterministic BQP to compute $\beta(R)$ as a mixed integer linear program:

$$
\begin{aligned}
\max_{x\in\Re^n, X\in\mathcal{S}^n} \quad & \sum_{i\in[n]} R_{ii}x_i + \sum_{i<j:i,j\in[n]} 2R_{ij}X_{ij} & \\
\text{s.t.} \quad & \sum_{j\in[n]} A_{kj}x_j = b_k, & \forall k \in [m], \\
& \sum_{j<i:j\in[n]} A_{kj}X_{ji} + (A_{ki} - b_k)\, x_i + \sum_{j>i:j\in[n]} A_{kj}X_{ij} = 0, & \forall k \in [m], i \in [n], \\
& X_{ij} \le x_i, \quad X_{ij} \le x_j & \forall i < j : i, j \in [n], \\
& X_{ij} \ge x_i + x_j - 1, & \forall i < j : i, j \in [n], \\
& x_i, X_{ij} \in \{0, 1\}, & \forall i, j \in [n].
\end{aligned}
\tag{24}
$$

The second set of constraints in (24) is obtained by multiplying each equality constraint in (23) with the variable x_i and then replacing each term $x_i x_j$ with X_{ij} and replacing x_i^2 with x_i for all i. While these additional set of mn linear equalities are redundant for the integer program, they are known to help tighten the linear programming relaxation of (24). Relaxing the integrality constraints to $x_i, X_{ij} \in [0,1]$, we obtain an upper bound on the optimal value of the deterministic BQP by solving a linear program formulated as follows:

$$\begin{aligned}
\beta_1(R) = \max_{x\in\Re^n, X\in\mathcal{S}^n} \quad & \sum_{i\in[n]} R_{ii}x_i + \sum_{i<j:i,j\in[n]} 2R_{ij}X_{ij} \\
\text{s.t.} \quad & \text{same set of constraints as in (24) except} \\
& \text{that the binary constraints are replaced by} \\
& 0 \le x_i \le 1, \qquad \forall i \in [n], \\
& X_{ij} \ge 0, \qquad \forall i<j : i,j \in [n],
\end{aligned} \tag{25}$$

where $\beta_1(R) \ge \beta(R)$. Taking the linear programming dual of (25) and plugging into (10), we obtain an upper bound on the expected optimal value by solving the convex optimization problem:

$$\begin{aligned}
\bar{\phi}_1 = \min_{v,w,\delta,\gamma,\lambda,R\in\mathcal{S}^n} \quad & \sum_{k\in[m]} b_k v_k + \sum_{i<j:i,j\in[n]} \gamma_{ij} + \sum_{i\in[n]} \lambda_i + \sum_{i\in[n]}\sum_{j\in[n]} E_{\theta_{ij}}(\tilde{Q}_{ij} - R_{ij})^+ \\
\text{s.t.} \quad & \sum_{k\in[m]} A_{ki}v_k + \sum_{k\in[m]} (A_{ki} - b_k)\, w_{ki} \\
& - \sum_{j\ne i:j\in[n]} \delta_{ij} + \sum_{i<j:j\in[n]} \gamma_{ij} + \sum_{j<i:j\in[n]} \gamma_{ji} + \lambda_i \ge R_{ii}, \qquad \forall i \in [n], \\
& \sum_{k\in[m]} A_{kj}w_{ki} + \sum_{k\in[m]} A_{ki}w_{kj} + \delta_{ij} + \delta_{ji} - \gamma_{ij} \ge 2R_{ij}, \qquad \forall i<j : i,j\in[n], \\
& \delta_{ij}, \delta_{ji}, \gamma_{ij} \ge 0, \qquad \forall i<j : i,j\in[n], \\
& \lambda_i \ge 0, \qquad \forall i \in [n],
\end{aligned}$$

where v, w, δ, γ, λ are the dual variables for each set of constraints in (25) in order. Clearly, $\bar{\phi}_1 \ge \phi$.

We can also develop semidefinite programming based bounds on ϕ using SDP relaxations for discrete optimization problems as developed in Körner [26], Shor [38] and Poljak, Rendl and Wolkowicz [33] among others. A semidefinite relaxation for BQP is given as:

$$\begin{aligned}
\beta_2(R) = \max_{x\in\Re^n, X\in\mathcal{S}^n} \quad & \sum_{i\in[n]}\sum_{j\in[n]} R_{ij}X_{ij} \\
\text{s.t.} \quad & \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succeq 0, \\
& \sum_{j\in[n]} A_{kj}x_j = b_k, \qquad \forall k \in [m], \\
& \sum_{j\in[n]} A_{kj}X_{ij} = b_k x_i, \qquad \forall k \in [m], i \in [n], \\
& X_{ii} = x_i, \qquad \forall i \in [n].
\end{aligned} \tag{26}$$

Note that $\beta_2(R)$ is computable in polynomial time and provides an upper bound on $\beta(R)$. We then obtain an upper bound on ϕ by solving the SDP:

$$\begin{aligned}
\bar{\phi}_2 = \min_{R,r,u,V,w} \quad & r + \sum_{i\in[n]}\sum_{j\in[n]} E_{\theta_{ij}}(\tilde{Q}_{ij} - R_{ij})^+ \\
\text{s.t.} \quad & \begin{bmatrix} r - b^T w & (A^T w - V^T b - u)^T/2 \\ (A^T w - V^T b - u)/2 & Diag(u) - R + \frac{1}{2}(A^T V + V^T A) \end{bmatrix} \succeq 0.
\end{aligned}$$

Clearly, $\bar{\phi}_2 \ge \phi$.

5.2. Approximations guarantee for ϕ. In this section, we show that the polynomial time computable upper bounds on ϕ obtained through the linear and semidefinite programming relaxations preserve the approximation guarantee of the deterministic BQPs if available. This brings us to the following result on the approximation guarantee on ϕ.

THEOREM 9. *Assume that the following condition holds: For all symmetric matrices Q in the set Ω, the optimal objective of a deterministic BQP is approximable within a factor of $(1-\alpha)$ for some $\alpha \in [0,1)$ by a polynomial time computable upper bound $\bar{\beta}(Q)$:*

$$(1-\alpha)\bar{\beta}(Q) \leq \beta(Q) \leq \bar{\beta}(Q). \tag{27}$$

Then, the bound ϕ on the expected optimal value of the random BQP in (4) is approximable within a factor of $(1-\alpha)$ by $\bar{\phi}$:

$$(1-\alpha)\bar{\phi} \leq \phi \leq \bar{\phi}, \tag{28}$$

where the upper bound $\bar{\phi}$ is defined as:

$$\bar{\phi} = \inf_{R \in \mathcal{S}^n, R \in \Omega} \left[\bar{\beta}(R) + \sum_{i \in [n]} \sum_{j \in [n]} E_{\theta_{ij}} (\tilde{Q}_{ij} - R_{ij})^+ \right]. \tag{29}$$

Proof. Note that it is easy to verify that in formulation (10) for the bound ϕ, we can restrict each optimal R_{ij} to lie in the support Ω_{ij}. It is clear that $\phi \leq \bar{\phi}$ by replacing the maximization problem $\beta(R) = \max_{x \in \mathcal{X}} x^T R x$ in formulation (10) with the upper bound $\bar{\beta}(R)$. To show that $\phi \geq (1-\alpha)\bar{\phi}$, observe that:

$$\begin{aligned}
\phi &\geq \inf_{R \in \mathcal{S}^n, R \in \Omega} \left[(1-\alpha)\bar{\beta}(R) + \sum_{i \in [n]} \sum_{j \in [n]} E_{\theta_{ij}} (\tilde{Q}_{ij} - R_{ij})^+ \right], \\
&= (1-\alpha) \inf_{R \in \mathcal{S}^n, R \in \Omega} \left[\bar{\beta}(R) + \frac{1}{1-\alpha} \sum_{i \in [n]} \sum_{j \in [n]} E_{\theta_{ij}} (\tilde{Q}_{ij} - R_{ij})^+ \right], \\
&\geq (1-\alpha) \inf_{R \in \mathcal{S}^n, R \in \Omega} \left[\bar{\beta}(R) + \sum_{i \in [n]} \sum_{j \in [n]} E_{\theta_{ij}} (\tilde{Q}_{ij} - R_{ij})^+ \right], \\
&= (1-\alpha)\bar{\phi},
\end{aligned}$$

where the first inequality is obtained from the relationship $\beta(R) \geq (1-\alpha)\bar{\beta}(R)$ and the second inequality is obtained from the nonnegativity of the term $E_{\theta_{ij}}(\tilde{Q}_{ij} - R_{ij})^+$ and $\alpha \in [0,1)$. ∎

This result helps extend the approximation bound for the deterministic BQP to the random BQP. For example, a classical result by Goemans and Williamson [19] provides a 0.878 approximation algorithm for the maximum cut (MAX-CUT) problem using a semidefinite program. Given a graph with a vertex set $[n]$ and nonnegative coefficients $Q_{ij} = Q_{ji}$ for each pair of vertices i and j, the maximum cut problem is to find a set of vertices $\mathcal{S} \subset [n]$ that maximizes the weight of the edges with one endpoint in $\mathcal{S}$ and the other in $[n] \setminus \mathcal{S}$. This is formulated as a integer quadratic program:

$$\begin{aligned}
\beta_{\text{max-cut}}(Q) \quad = \quad \max \quad & \frac{1}{4} \sum_{i \neq j: i,j \in [n]} Q_{ij} (1 - x_i x_j) \\
\text{s.t.} \quad & x_i \in \{-1, 1\}, \qquad \forall i \in [n],
\end{aligned} \tag{30}$$

where $x_i = 1$ if vertex i is on one side of the cut (say set $\mathcal{S}$) and $x_i = -1$ if the vertex is on the other side of the cut (set $[n] \setminus \mathcal{S}$). Goemans and Williamson [19] proposed solving the following semidefinite relaxation:

$$(31)\qquad \begin{aligned} \bar{\beta}_{\text{sdp-max-cut}}(Q) \quad = \quad \max \quad & \frac{1}{4}\sum_{i \neq j: i,j \in [n]} Q_{ij}\left(1 - X_{ij}\right) & \\ \text{s.t.} \quad & X_{ii} = 1, & \forall i \in [n], \\ & X \succeq 0, & \end{aligned}$$

where the rank one constraint on the matrix X is dropped. While formulation (30) is NP-hard to solve, the semidefinite relaxation in (31) is solvable in polynomial time. They showed that the semidefinite program provides an 0.878 approximation for the max-cut problem:

$$(32)\qquad 0.878\bar{\beta}_{\text{sdp-max-cut}}(Q) \leq \beta_{\text{max-cut}}(Q) \leq \bar{\beta}_{\text{sdp-max-cut}}(Q),$$

for nonnegative matrices Q. Consider a random max-cut problem where the weights of the edges are nonnegative random variables and we are interested in computing the bound:

$$(33)\qquad \phi_{\text{max-cut}} \quad = \quad \sup_{\theta \in \Theta} E_\theta\left[\beta_{\text{max-cut}}(\tilde{Q})\right].$$

An upper bound on $\phi_{\text{max-cut}}$ is then obtained by using the semidefinite programming relaxation of the max-cut problem in (10) as follows:

(34)

$$\bar{\phi}_{\text{sdp-max-cut}} = \inf_{R \geq 0}\left[\max_{X_{ii}=1, \forall i, X \succeq 0} \frac{1}{4}\sum_{i \neq j} R_{ij}\left(1 - X_{ij}\right) + \sum_{i,j} E_{\theta_{ij}}\left(\tilde{Q}_{ij} - R_{ij}\right)^+\right].$$

Define the entries of a Laplacian matrix of the graph weighted with the matrix R as follows:

$$L_{ij}(R) = \begin{cases} \sum_k R_{ik}, & \text{if } i = j, \\ -R_{ij}, & \text{if } i \neq j. \end{cases}$$

We can then reformulate the inner semidefinite program in (34) as follows:

$$\begin{aligned} \bar{\beta}_{\text{sdp-max-cut}}(R) \quad = \quad \max \quad & \frac{1}{4}L(R) \cdot X & \\ \text{s.t.} \quad & X_{ii} = 1, & \forall i \in [n], \\ & X \succeq 0, & \end{aligned}$$

where it is easy to verify that strong duality holds for the semidefinite relaxation of the max-cut problem. Taking the dual of this semidefinite program and plugging back into (34), we obtain the polynomial time computable upper bound:

$$(35)\qquad \bar{\phi}_{\text{sdp-max-cut}} = \inf_{R \geq 0, Diag(y) \succeq L(R)}\left[\frac{1}{4}\sum_i y_{ii} + \sum_{i,j} E_{\theta_{ij}}\left(\tilde{Q}_{ij} - R_{ij}\right)^+\right].$$

Applying the result from Theorem 9, we obtain the following result.

Corollary 10. *Assume that the entries of the matrix $\tilde{Q}$ are nonnegative random variables. Then the semidefinite programming bound $\bar{\phi}_{sdp\text{-}max\text{-}cut}$ provides an 0.878 approximation of $\phi_{max\text{-}cut}$ as follows:*

$$(36)\qquad 0.878\bar{\phi}_{sdp\text{-}max\text{-}cut} \leq \phi_{max\text{-}cut} \leq \bar{\phi}_{sdp\text{-}max\text{-}cut}.$$

5.3. A subgradient method to compute ϕ. To numerically compute the bound ϕ in (10), we propose a subgradient method in this section where BQPs are solved at each iteration of the algorithm. Let us denote the objective function of the minimization problem as follows:

$$f(R) = \max_{x \in \mathcal{X}} x^T R x + \sum_{i \in [n]} \sum_{j \in [n]} E_{\theta_{ij}} (\tilde{Q}_{ij} - R_{ij})^+. \tag{37}$$

A subgradient method to minimize $f(R)$ is provided in Algorithm 1.

Algorithm 1 A subgradient algorithm to compute ϕ

Define $R^{(0)}$ with $f_{best} = f(R^{(0)})$ and $R_{best} = R^{(0)}$.
for $k = 1, 2, 3, \ldots$ **do**
  Update $R^{(k)} = R^{(k-1)} - \alpha_k\, g^{(k-1)}$,
  where $g^{(k-1)}$ is a subgradient of $f(\cdot)$ at $R^{(k-1)}$, and α_k is the step size.
  if $f(R^{(k)}) < f_{best}$ **then**
    Update $f_{best} = f(R^{(k)})$, and $R_{best} = R^{(k)}$.
  end if
end for

In general, the subgradient method is known to converge slowly. Moreover, for our problem we must solve a BQP at every iteration making the computation of ϕ hard in general. Let $f_{best}^{(k)}$ denote the best objective value obtained after k iterations using the subgradient method. Let R^* be a minimizer of $f(R)$, and let f^* denote the minimum value. With $\|R^{(0)} - R^*\|_2 = \epsilon_0$ and the norm of the subgradient $\|g^{(k)}\|_2 \le G$, a basic inequality for the subgradient method (see Boyd et al. [10]) is:

$$f_{best}^{(k)} - f^* \le \frac{\epsilon_0^2 + G^2 \sum_{i \in [k]} \alpha_i^2}{2 \sum_{i \in [k]} \alpha_i} \tag{38}$$

where α_i is the step size taken at the ith iteration. In the numerical tests, suppose we terminate Algorithm 1 after k steps, then we can select the step size α_i to minimize the right hand of the basic inequality (38). The optimal occurs when all α_i are equal to a constant $\alpha_{opt} = \frac{\epsilon_0}{G\sqrt{k}}$. In our problem, all the components of the subgradient g are in $[-1, 1]$ and hence $\|g\|_2 \le n$. Hence we can set $G = n$. With the step size α_{opt}, we have:

$$f_{best}^{(k)} - f^* \le \frac{n\epsilon_0}{\sqrt{k}}.$$

However, since one typically does not know the value of ϵ_0, in our numerical experiments we set the step size to:

$$\alpha_i = \alpha := \frac{1}{n\sqrt{k}}, \quad i = 1, \ldots, k.$$

With the above choice of step sizes, it can be guaranteed that:

$$f_{best}^{(k)} - f^* \le \frac{n(\epsilon_0^2 + 1)}{2\sqrt{k}}. \tag{39}$$

Therefore, using the constant step size α, to find an ϵ-optimal solution to (37), we need $k = \frac{n^2(\epsilon_0^2+1)^2}{4\epsilon^2}$ iterations.

5.4. Numerical experiments. In this section, we conduct numerical experiments to test the strength of the linear programming and semidefinite programming based bounds for random QUBO and quadratic knapsack problems. The computational studies were implemented in Matlab R2014b on an Intel Core i7-CPU (2.6GHz) laptop with 4GB of RAM. The SDP problems were solved with SDPT3 version 4 ([40, 41]) using CVX ([21, 20]) and the BQPs were solved with CPLEX 12.6 using the Matlab interface. We consider the setting where only the mean and variance of the entries for $\tilde{Q}$ are given. We compute the gap between the LP and SDP bounds $\bar{\phi}_1$ and $\bar{\phi}_2$ and ϕ. The relative gap between the bounds is denoted by $gap_1 = \frac{\bar{\phi}_1 - \phi}{\phi}$ and $gap_2 = \frac{\bar{\phi}_2 - \phi}{\phi}$ respectively. To estimate ϕ, we use the subgradient method as described in the previous section. Let ϕ_{num} denote the numerical value obtained by the subgradient method. From (39), we know that a lower bound on ϕ after k iterations is:

$$\underline{\phi} = \phi_{num} - \frac{n(\epsilon_0^2 + 1)}{2\sqrt{k}}. \tag{40}$$

Define $\underline{gap}_1 = \frac{\bar{\phi}_1 - \phi_{num}}{\phi_{num}}$ and $\overline{gap}_1 = \frac{\bar{\phi}_1 - \phi}{\underline{\phi}}$ for the LP bound. Since $\underline{\phi} \le \phi \le \phi_{num}$, we have $\underline{gap}_1 \le gap_1 \le \overline{gap}_1$. Similarly, we define $\underline{gap}_2$ and $\overline{gap}_2$ for the SDP bound. In our computations, we use $10000n^2/\bar{\phi}_2^2$ iterations to compute ϕ_{num}. Lastly, since we do not know the exact value of ϵ_0, we estimate this value by $\hat{\epsilon}_0 = \|R^{(0)} - R_{best}\|_2$, where R_{best} is the numerical solution obtained from the subgradient method. The CPU time (in seconds) of computing the LP bound $\bar{\phi}_1$ and SDP bound $\bar{\phi}_2$ are denoted by $t(\bar{\phi}_1)$ and $t(\bar{\phi}_2)$ while the total computational time to solve the BQPs and applying the subgradient method for computing the numerical approximation ϕ_{num} of the bound ϕ is denoted by $t(\phi_{num})$.

5.4.1. QUBO instances. We generate the mean value for each entry of $\tilde{Q}$ from a standard normal distribution, and the standard deviation for each entry of $\tilde{Q}$ from a standard uniform distribution independently. Thus $\mu_{ij} \sim N(0,1)$, $\sigma_{ij} \sim U(0,1)$, $i, j \in [n], i \le j$ and $\mu_{ji} = \mu_{ij}, \sigma_{ji} = \sigma_{ij}$. For a given symmetric matrix of size $n \times n$, we generate 10 instances by randomly sampling the mean and variance from this distribution. The numerical results for the QUBO instances with $n = 10$ and $n = 50$ are shown in Table 3 and 4 respectively[1]. From the numerical results, we see that the SDP upper bound is significantly tighter than the LP bound, particularly for the larger instances and they are fairly close to the numerical approximation of the bound ϕ. The computational times to solve the LP and SDP bounds are also significantly smaller than the time required for the subgradient method as expected.

Table 3: QUBO, $n = 10$

Instance	$\bar{\phi}_1$	$\bar{\phi}_2$	ϕ_{num}	$\underline{\phi}$	$\hat{\epsilon}_0$	$gap_1(\%)$	$gap_2(\%)$	$t(\bar{\phi}_1)$	$t(\bar{\phi}_2)$	$t(\phi_{num})$
1	34.94	32.67	31.68	31.44	0.614	[10.29,11.13]	[3.13,3.91]	0.20	0.30	308.83
2	31.85	29.10	28.46	28.26	0.576	[11.91,12.70]	[2.25,2.97]	0.17	0.20	397.14
3	36.68	33.13	32.27	31.96	0.870	[13.67,14.77]	[2.67,3.66]	0.20	0.30	307.71
4	40.14	39.41	38.16	37.87	0.623	[5.19,5.99]	[3.28,4.07]	0.22	0.33	217.22
5	49.91	48.72	47.40	47.07	0.528	[5.30,6.03]	[2.79,3.51]	0.17	0.25	143.48
6	31.33	29.98	28.87	28.59	0.874	[8.52,9.58]	[3.84,4.86]	0.22	0.36	367.37
7	39.21	35.11	33.77	33.47	0.783	[16.11,17.15]	[3.97,4.90]	0.17	0.22	268.27
8	49.55	47.40	45.72	45.39	0.574	[8.38,9.17]	[3.67,4.43]	0.20	0.30	150.12
9	57.41	56.23	55.28	54.90	0.527	[3.85,4.57]	[1.72,2.42]	0.22	0.22	108.58
10	29.39	27.83	26.77	26.56	0.665	[9.79,10.66]	[3.96,4.78]	0.20	0.22	439.08

[1] All test instances can be obtained from the first author.

Table 4: QUBO, $n = 50$

Instance	$\bar{\phi}_1$	$\bar{\phi}_2$	ϕ_{num}	$\underline{\phi}$	$\hat{\epsilon}_0$	$gap_1(\%)$	$gap_2(\%)$	$t(\bar{\phi}_1)$	$t(\bar{\phi}_2)$	$t(\phi_{num})$
1	931.11	725.16	709.35	705.29	0.339	[31.26,32.02]	[2.23,2.82]	1.00	6.00	779.25
2	983.86	799.72	781.81	776.97	0.461	[25.84,26.63]	[2.29,2.93]	0.50	4.50	379.62
3	955.62	736.83	716.21	711.66	0.486	[33.43,34.28]	[2.88,3.54]	0.90	7.00	436.52
4	953.41	747.08	730.87	726.09	0.520	[30.45,31.31]	[2.22,2.89]	1.05	6.90	406.91
5	941.52	720.87	701.62	697.41	0.411	[34.19,35.00]	[2.74,3.36]	0.40	4.80	812.48
6	997.60	791.47	773.96	769.28	0.413	[28.90,29.68]	[2.26,2.88]	0.90	6.10	345.85
7	924.80	726.51	708.98	704.07	0.590	[30.44,31.35]	[2.47,3.19]	0.50	6.90	541.82
8	951.17	733.24	716.46	712.00	0.461	[32.76,33.59]	[2.34,2.98]	0.45	5.50	501.41
9	972.19	779.59	760.89	756.26	0.433	[27.77,28.55]	[2.46,3.09]	0.60	8.00	495.07
10	958.79	762.18	743.88	739.11	0.502	[28.89,29.72]	[2.46,3.12]	1.00	4.80	492.61

5.4.2. QKP instances. The quadratic knapsack problem was introduced by Gallo et al. [16]. Assume that n items are given where item i has a positive weight w_i. In addition we are given a profit matrix Q, where Q_{ii} is the profit achieved if item i is selected and $Q_{ij}+Q_{ji}$ is the profit achieved if both item i and item j are selected. The purpose is to select a set of items whose total weight does not exceed a given knapsack capacity b, so as to maximize the overall profit. The problem is formulated as

$$\beta_{\text{qkp}}(Q) \quad = \quad \max \left\{ \sum_{i\in[n]} \sum_{j\in[n]} Q_{ij} x_i x_j \;\middle|\; \sum_{i\in[n]} w_i x_i \leq b, x_i \in \{0,1\}, \forall i \in [n] \right\}.$$

Assume the profit matrix is a symmetric random matrix denoted by $\tilde{Q}$ and the mean and variance of $\tilde{Q}_{ij}$ are given. The data are randomly generated as follows: Each weight w_i is a randomly generated integer in $[1, 50]$, and the capacity b is randomly distributed in $[\min_{i\in[n]}\{w_i\}, \sum_{i\in[n]} w_i]$. The mean and and standard deviation of the profit $\tilde{Q}_{ij}$ are generated as $\mu_{ij} \sim U(0, 10)$, $\sigma_{ij} \sim U(0, 20)$, $i, j \in [n], i \leq j$, and $\mu_{ji} = \mu_{ij}, \sigma_{ji} = \sigma_{ij}$. We generate 10 random instances for a given n. The numerical results for $n = 20$ and $n = 40$ are shown in Tables 5 and Table 6. From the numerical results, we see that the LP bound is closer to the SDP bound for the QKP instances and both the bounds are quite close to the numerical approximation of ϕ. Moreover, they can typically be computed at a fraction of the time for computing ϕ_{num}.

Table 5: QKP, $n = 20$

Instance	$\bar{\phi}_1$	$\bar{\phi}_2$	ϕ_{num}	$\underline{\phi}$	$\hat{\epsilon}_0$	$gap_1(\%)$	$gap_2(\%)$	$t(\bar{\phi}_1)$	$t(\bar{\phi}_2)$	$t(\phi_{num})$
1	75.35	75.35	73.54	72.54	1.262	[2.46,3.87]	[2.46,3.87]	0.60	1.00	332.92
2	127.41	127.10	125.75	124.71	0.774	[1.32,2.17]	[1.07,1.92]	0.22	0.36	178.28
3	114.47	114.33	113.13	112.27	0.697	[1.18,1.95]	[1.06,1.83]	0.35	0.45	204.99
4	195.29	194.30	192.70	191.37	0.585	[1.34,2.05]	[0.83,1.53]	0.30	0.40	227.40
5	204.48	202.64	201.64	200.41	0.436	[1.41,2.03]	[0.50,1.11]	0.30	0.45	424.53
6	258.05	257.13	255.24	253.67	0.444	[1.10,1.73]	[0.74,1.36]	0.45	0.95	168.45
7	97.24	97.42	95.64	94.79	0.842	[1.66,2.57]	[1.86,2.77]	0.35	0.45	207.50
8	331.11	328.29	327.50	325.79	0.179	[1.10,1.63]	[0.24,0.77]	0.30	0.45	23.06
9	83.92	83.91	82.73	81.91	0.954	[1.44,2.45]	[1.43,2.44]	0.35	0.75	300.69
10	41.73	41.73	40.57	39.89	1.469	[2.86,4.61]	[2.86,4.61]	0.60	.130	1112.64

Table 6: QKP, $n = 40$

Instance	$\bar{\phi}_1$	$\bar{\phi}_2$	ϕ_{num}	$\underline{\phi}$	$\hat{\epsilon}_0$	$gap_1(\%)$	$gap_2(\%)$	$t(\bar{\phi}_1)$	$t(\bar{\phi}_2)$	$t(\phi_{num})$
1	1144	1133.92	1131.07	1125.16	0.232	[1.14,1.67]	[0.25,0.78]	1.70	3.40	8452.63
2	382.22	379.61	377.62	375.50	0.333	[1.22,1.79]	[0.53,1.09]	1.30	3.60	19761.31
3	1345.70	1330.11	1328.49	1321.76	0.225	[1.30,1.81]	[0.12,0.63]	1.05	2.09	38.39
4	386.28	384.24	382.59	380.45	0.327	[0.96,1.53]	[0.43,1.00]	1.25	5.50	14026.03
5	748.67	739.46	736.40	732.63	0.138	[1.67,2.19]	[0.42,0.93]	2.10	3.20	18706.46
6	1261.12	1254.80	1252.66	1246.56	0.000	[0.67,1.17]	[0.17,0.66]	1.00	3.00	7252.90
7	639.89	633.04	630.46	627.13	0.199	[1.50,2.03]	[0.41,0.94]	0.90	2.90	24719.21
8	1127.60	1118.89	1115.93	1109.88	0.280	[1.05,1.60]	[0.27,0.81]	1.00	3.40	8460.81
9	1333.53	1316.79	1314.88	1308.36	0.138	[1.42,1.92]	[0.14,0.64]	1.20	3.60	38.69
10	1290.90	1283.80	1281.62	1274.79	0.260	[0.72,1.26]	[0.17,0.71]	1.05	4.00	370.50

6. Conclusions. In this paper, we have developed a new bound on the expected optimal value of a random binary quadratic program. The bound makes use of only marginal distributional information of the objective coefficients and is thus robust to all dependencies among the coefficients. We show that the bound has several attractive properties. Specifically it preserves polynomial time computability when the instance of the binary quadratic program is easy. It also preserves approximation guarantees of linear and semidefinite programming relaxations for the NP hard instances. We also develop new closed form bounds for the quadratic assignment and quadratic k-cluster problem that are valid even when the assumption of independence is violated.

To the best of our knowledge, this is one of the few works that looks at the expected optimal value of a binary quadratic program and we hope this will garner more interest from the research community. Some of the possible future research directions in this topic includes incorporating additional distributional information to tighten the bounds, developing stronger approximations for the bound ϕ by using valid inequalities, developing new random test instances from the distribution that attains the bound ϕ and possibly developing randomized solutions for the random version of the problem. We leave these topics for future research.

Acknowledgements. The authors would like to thanks the Associate Editor Samuel Burer and the review team for providing useful inputs and comments that helped improve the paper.

REFERENCES

[1] D. Aldous, *The ζ(2) limit in the random assignment problem*, Random Structures and Algorithms, 18 (2001), pp. 381–418.

[2] B. Alidaee, G. A. Kochenberger, and A. Ahmadian, *0-1 quadratic programming approach for optimum solutions of two scheduling problems*, International Journal of Systems Science, 25 (1994), pp. 401–408.

[3] F. Barahona, M. Grötschel, M. Jünger, and G. Reinelt, *An application of combinatorial optimization to statistical physics and circuit layout design*, Operations Research, 36 (1988), pp. 493–513.

[4] J. H. Beardwood, J. Halton and J. M. Hammersley, *The shortest path through many points*, Mathematical Proceedings of the Cambridge Philosophical Society, 55 (1959), pp. 299–327.

[5] J. E. Beasley, *Heuristic algorithms for the unconstrained binary quadratic programming problem*, London, UK: Management School, Imperial College, 4 (1998).

[6] D. Bertsimas, K. Natarajan, and C. P. Teo, *Probabilistic combinatorial optimization: Moments, semidefinite programming, and asymptotic bounds*, SIAM Journal on Optimization, 15 (2004), pp. 185–209.

[7] D. Bertsimas, K. Natarajan, and C. P. Teo, *Persistence in discrete optimization under data uncertainty*, Mathematical Programming, 108 (2006), pp. 251–274.

[8] J. R. Birge and F. Louveaux, *Introduction to stochastic programming*, Springer Series in Operations Research and Financial Engineering, 2 ed., 2011.

[9] J. R. Birge and M. J. Maddox, *Bounds on expected project tardiness*, Operations Research, 43 (1995), pp. 838–850.

[10] S. Boyd, L. Xiao, and A. Mutapcic, *Subgradient methods*, lecture notes of EE392o, Stanford University, Autumn Quarter, 2004 (2003).
[11] R. E. Burkard and U. Fincke, *Probabilistic asymptotic properties of some combinatorial optimization problems*, Discrete Applied Mathematics, 12 (1985), pp. 21–29.
[12] M. W. Carter, *The indefinite zero-one quadratic problem*, Discrete Applied Mathematics, 7 (1984), pp. 23–44.
[13] J. Cheng, E. Delage, and A. Lisser, *Distributionally robust stochastic knapsack problem*, SIAM Journal of Optimization, 24 (2014), pp. 1485–1506.
[14] V. H. M. V. R. K. A. H. G. Frenk, J. B. B., *Asymptotic properties of the quadratic assignment problem*, Mathematics of Operations Research, 10 (1985), pp. 100–116.
[15] A. M. Frieze, *On the value of a random minimum spanning tree problem*, Discrete Applied Mathematics, 10 (1985), pp. 47–56.
[16] G. Gallo, P. L. Hammer, and B. Simeone, *Quadratic knapsack problems*, in Combinatorial Optimization, Springer, 1980, pp. 132–149.
[17] B. Ghaddar, J. C. Vera, and M. F. Anjos, *Second-order cone relaxations for binary quadratic polynomial programs*, SIAM Journal of Optimization, 21 (2011), pp. 391–414.
[18] F. Glover, G. A. Kochenberger, and B. Alidaee, *Adaptive memory search for binary quadratic programs*, Management Science, 44 (1998), pp. 336–345.
[19] M. X. Goemans and D. P. Williamson, *Improved approximation algorithms for max cut and satisability problems using semidefinite programming*, Journal of ACM, 42.
[20] M. Grant and S. Boyd, *Graph implementations for nonsmooth convex programs*, Recent Advances in Learning and Control (a tribute to M. Vidyasagar), V. Blondel, S. Boyd, and H. Kimura, editors, Lecture Notes in Control and Information Sciences, (2008), pp. 95–110.
[21] M. Grant and S. Boyd, *Cvx research, inc. cvx: Matlab software for disciplined convex programming, version 2.0. beta.*, (2012), http://cvxr.com/cvx.
[22] J. N. Hagstrom, *Computational complexity of pert problems*, Networks, 18 (1988), pp. 139–147.
[23] C. Helmberg and F. Rendl, *Solving quadratic (0, 1)-problems by semidefinite programs and cutting planes*, Mathematical Programming, 82 (1998), pp. 291–315.
[24] M. Jünger, A. Martin, G. Reinelt, and R. Weismantel, *Quadratic 0/1 optimization and a decomposition approach for the placement of electronic circuits*, Mathematical Programming, 63 (1994), pp. 257–279.
[25] S. Kim and M. Kojima, *Second order cone programming relaxation of nonconvex quadratic optimization problems*, Optimization Methods and Software, 15 (2001), pp. 201–224.
[26] F. Körner, *A tight bound for the boolean quadratic optimization problem and its use in a branch and bound algorithm*, Optimization, 19 (1988), pp. 711–721.
[27] J. B. Lasserre, *A max-cut formulation of 0/1 programs*, Operations Research Letters, 44 (2016), pp. 158–164.
[28] D. J. Laughhunn, *Quadratic binary programming with application to capital-budgeting problems*, Operations Research, 18 (1970), pp. 454–461.
[29] A. J. McNeil, R. Frey, and P. Embrechts, in Quantitative Risk Management: Concepts, Techniques and Tools, Princeton University Press, 2005.
[30] I. Meilijson and A. Nadas, *Convex majorization with an application to the length of critical path*, Journal of Applied Probability, 16 (1979), pp. 671–677.
[31] P. M. Pardalos and G. P. Rodgers, *Computational aspects of a branch and bound algorithm for quadratic zero-one programming*, Computing, 45 (1990), pp. 131–144.
[32] J. C. Picard and H. D. Ratliff, *Minimum cuts and related problems*, Networks, 5 (1975), pp. 357–370.
[33] S. Poljak, F. Rendl, and H. Wolkowicz, *A recipe for semidefinite relaxation for (0, 1)-quadratic programming*, Journal of Global Optimization, 7 (1995), pp. 51–73.
[34] W. T. Rhee, *Stochastic analysis of the quadratic assignment problem*, Mathematics of Operations Research, 16 (1991), pp. 223–239.
[35] C. Rother, V. Kolmogorov, V. Lempitsky, and M. Szummer, *Optimizing binary mrfs via extended roof duality*, IEEE Conference on Computer Vision and Pattern Recognition, (2007), pp. 1–8.
[36] H. Scarf, *A min-max solution of an inventory problem*, Studies in The Mathematical Theory of Inventory and Production, (1958), pp. 201–209.
[37] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming: Modeling and theory*, MOS-SIAM Series on Optimization, (2009).
[38] N. Z. Shor, *Class of global minimum bounds of polynomial functions*, Cybernetics and Systems Analysis, 23 (1987), pp. 731–734.
[39] W. Szpankowski, *Combinatorial optimization problems for which almost every algorithm in*

asymptotically optimal, Optimization, 33 (1995), pp. 359–367.

[40] K. C. Toh, M. J. Todd, and R. H. Tutuncu, *SDPT3 — a matlab software package for semidefinite programming*, Optimization Methods and Software, 11 (1999), pp. 545–581.

[41] K. C. Toh, M. J. Todd, and R. H. Tutuncu, *Solving semidefinite-quadratic-linear programs using SDPT3*, Mathematical Programming, 95 (2003), pp. 189–217.

[42] G. Weiss, *Stochastic bounds on distributions of optimal value functions with applications to PERT, network flows and reliability*, Operations Research, 34 (1986), pp. 595–605.