Poster Abstracts

1 Bugra Can, Rutgers University, U.S.A.

A Variance-Reduced Stochastic Accelerated Primal Dual Algorithm

In this work, we consider strongly convex strongly concave (SCSC) saddle point (SP) problems $\min_{x\in\mathbb{R}^{d_x}} \max_{y\in\mathbb{R}^{d_y}} f(x,y)$ where f is L-smooth, $f(\cdot,y)$ is μ-strongly convex for every y, and $f(x,\cdot)$ is μ-strongly concave for every x. Such problems arise frequently in machine learning in the context of robust empirical risk minimization (ERM), e.g., *distributionally robust* ERM, where partial gradients are estimated using mini-batches of data points. Assuming we have access to an unbiased stochastic first-order oracle, we consider the stochastic accelerated primal dual (SAPD) algorithm recently introduced in [Zhang et al, 2021] for SCSC SP problems as a robust method against the gradient noise. In particular, SAPD recovers the well-known stochastic gradient descent ascent (SGDA) as a special case when the momentum parameter is set to zero and can achieve an accelerated rate when the momentum parameter is properly tuned, i.e., improving the $\kappa \triangleq L/\mu$ dependence from κ^2 for SGDA to κ. We propose efficient variance-reduction strategies for SAPD based on Richardson-Romberg extrapolation and show that our method improves upon SAPD both in practice and in theory.

2 Sebastien Colla, Université Catholique de Louvain, Belgium

Optimization of Reactive Ink Formulation for Additive Manufacturing of Charged Membranes

We develop a methodology to automatically compute worst-case performance bounds for a large class of first-order decentralized optimization algorithms. These algorithms aim at minimizing the average of local functions that are distributed across a network of agents. They typically combine local computations and consensus steps. Our methodology is based on the approach of Performance Estimation Problem (PEP), which allow computing the worst-case performance and worst-case instance of first-order optimization algorithms by solving an SDP. We propose a PEP formulation that can embed averaging consensus steps and provides guarantees valid over an entire class of averaging matrices, characterized by their spectral range. This formulation often allows recovering the worst possible averaging matrix for the given algorithm. We apply our methodology to different decentralized methods. For each of them, we obtain numerically tight worst-case performance bounds that significantly improve on the existing ones, as well as insights about the parameter tuning and the worst communication networks.

3 Nicole Cortes, University of Notre Dame, U.S.A.

Co-optimizing the Design and Operation strategy of solid oxide fuel cell-based hydrogen-electricity co-production systems

With renewable energy being the fastest growing electricity generation category [1] energy infrastructures and markets will become more volatile, requiring more flexible energy systems to balance supply and demand [2]. Integrated energy systems (IES) can offer this flexibility by exploiting synergies between multiple technologies. While optimization of IES concepts is present in current literature, detailed market analysis is a challenging task, calling for accurate, often non-linear, models of system operations as well as capturing the complex economic interactions that these markets involve. In this work, we present a framework for conducting optimization-based market-informed techno-economic analysis (TEA) of IES, formulated as a generalized disjunctive programming (GDP) model and implemented in Pyomo. Detailed equation-oriented process models are developed in the IDAES modeling platform [3-5]. We then use ALAMO to generate algebraic surrogates for operating costs, capital costs, and co-production constraints. Using these surrogates embedded in the GDP optimization model, we can directly compare the economic performance of different IES concepts. Here, we demonstrate the platform's capabilities by co-optimization of system design and operation of solid oxide fuel cell (SOFC)-based IES that coproduce electricity and hydrogen. We find that SOFC systems offer significant cost and technical advantages over alternatives, which is consistent with the finding of prior more traditional, e.g., levelized cost of electricity, TEA [6]. Moreover, the use of this framework demonstrates the importance of utilizing optimization models to drive technical and economic decision-making.

4 Niloofar Fadavi, Southern Methodist University, U.S.A.

An active-set method for two stage stochastic quadratic programming

In this poster, we examine two-stage stochastic quadratic programming problems. The first and second stage problems are convex programs with quadratic objective functions and affine constraints. We model the uncertainty that affects the right-hand sides and variable bounds in the second stage using random variables with finite support. We present exact and inexact proximal bundle methods to solve these problems. The exact method requires solving large-scale quadratic programs for each scenario in every iteration to generate the model of the expectation valued objective. To address the computational difficulty associated with this step, we utilize approximate second-stage solutions in the inexact variant of the algorithm. Using the relationship between the partitions generated from a primal-dual active-set method and the observations of random variables, we compute approximate solutions by solving (reduced) linear systems. We present the convergence and numerical analysis of these algorithms.

5 Jun-ya Gotoh, Chuo University, Japan
Knot Selection of B-Spline Regression via Trimmed Regularizer
B-spline regression is a nonlinear model estimation method that uses piecewise polynomial basis functions whose shape is determined by the given positions of knots. In this study, we propose a continuous optimization approach for accomplishing the selection of used knots (or bases) and the model fitting at the same time. More specifically, we introduce the cardinality constraint on the linearly-transformed coefficient vector of the model, reduce it to a partially-regularized least square problem, and apply GIST (Gong, et al. 2013). Numerical results show that the proposed method found better models than existing method.

6 Fadi Hamad, University of Pittsburgh, U.S.A.
A fully adaptive trust-region method
Adaptive second-order methods including ARC and TRACE attempt to maintain strong convergence guarantees without depending on conservative estimates of problem properties such as Lipschitz constants. However, on close inspection one can show these 'adaptive methods' have theoretical guarantees with severely suboptimal dependence on the Lipschitz constant of the Hessian. Our trust-region method, a simple modification of the classic trust-region method, finds first-order stationary points of functions that are bounded below and have a Lipschitz continuous Hessian, and has a convergence guarantee that is optimal with respect to all problem parameters up to an additive logarithmic term. On the CUTEst test set our method performs competitively compared with ARC and a classic trust-region method.

7 Yao Ji, School of Industrial Engineering, Purdue University, U.S.A.
Distributed Sparse Regression via Penalization
We study sparse linear regression over a network of agents, modeled as an undirected graph (with no centralized node). The estimation problem is formulated as the minimization of the sum of the local LASSO loss functions plus a quadratic penalty of the consensus constraint—the latter being instrumental to obtain distributed solution methods. While penalty-based consensus methods have been extensively studied in the optimization literature, their statistical and computational guarantees in the high dimensional setting remain unclear. This work provides an answer to this open problem. Our contribution is two-fold. First, we establish statistical consistency of the estimator: under a suitable choice of the penalty parameter, the optimal solution of the penalized problem achieves near optimal minimax rate $\mathcal{O}(s \log d/N)$ in ℓ_2-loss, where s is the sparsity value, d is the ambient dimension, and N is the total sample size in the network–this matches centralized sample rates. Second, we show that the proximal-gradient algorithm applied to the penalized problem, which naturally leads to distributed implementations, converges linearly up to a tolerance of the order of the centralized statistical error–the rate scales as $\mathcal{O}(d)$, revealing an unavoidable speed-accuracy dilemma. Numerical results demonstrate the tightness of the derived sample rate and convergence rate scalings. This presentation is based on the arXiv preprint "Distributed Sparse Regression via Penalization."

8 David Kiessling, KU Leuven, Belgium
Efficient Numerical Algorithms for Time Optimal Control
Our algorithms are motivated by Time Optimal Control Problems (TOCP) with nonlinear constraints which arise in motion planning of mechatronic systems. The resulting Nonlinear Programs (NLP) are obtained through direct multiple shooting discretization. We propose Sequential Linear Programming (SLP) algorithms that provide locally quadratic convergence in the case of fully determined problems. Due to the nonlinear constraints, our algorithms apply a trust-region globalization strategy ensuring global convergence. One variant of the algorithm keeps all iterates feasible enabling early termination at suboptimal, feasible solutions which is favorable for real-time applications. This additional feasibility is achieved by an efficient iterative strategy using evaluations of constraints, i.e., zero-order information. Convergence of the feasibility iterations can be enforced by reduction of the trust-region radius. These feasibility iterations maintain feasibility for general NLP. Therefore, our algorithms are applicable to general NLP. Our algorithms' efficiency is demonstrated for a TOCP motion planning simulation case of a mechatronic system.

9 J. Lyle Kim, Rice University, U.S.A.
Convergence and Stability of the Stochastic Proximal Point Algorithm with Momentum
Stochastic gradient descent with momentum

(SGDM) is the dominant algorithm in many optimization scenarios, including convex optimization instances and non-convex neural network training. Yet, in the stochastic setting, momentum interferes with gradient noise, often leading to specific step size and momentum choices in order to guarantee convergence, set aside acceleration. Proximal point methods, on the other hand, have gained much attention due to their numerical stability and elasticity against imperfect tuning. Their stochastic accelerated variants though have received limited attention how momentum interacts with the stability of (stochastic) proximal point methods remains largely unstudied. To address this, we focus on the convergence and stability of the stochastic proximal point algorithm with momentum (SPPAM), and show that SPPAM allows a faster linear convergence rate compared to stochastic proximal point algorithm (SPPA) with a better contraction factor, under proper hyperparameter tuning. In terms of stability, we show that SPPAM depends on problem constants more favorably than SGDM, allowing a wider range of step size and momentum that lead to convergence.

10 Clement Lezane, University of Twente, The Netherlands

Algorithms for Stochastic Complementary Composite Minimization

Inspired by regularization techniques in machine learning, we initiate the study of complementary composite minimization in the stochastic setting. This problem corresponds to the minimization of the sum of a smooth stochastic objective with a deterministic strongly convex (possibly non-smooth and non-Lipschitz) regularization term. Despite intensive work on closely related settings, prior to our work no complexity results for this problem were known. We close this gap by providing novel excess risk bounds, both for in-expectation as well as with high probability, guarantees.

11 Yongchun Li, Virginia Tech, U.S.A.

D-optimal Data Fusion: Exact and Approximation Algorithms

This paper studies the D-optimal Data Fusion (DDF) problem, which aims to select new data points considering the existing ones to maximize the logarithm of the determinant of overall Fisher information matrix. We show that the DDF problem is NP-hard and has no constant-factor polynomial-time approximation algorithm unless P = NP. Therefore, to solve the DDF problem effectively, we propose two equivalent convex integer programming formulations and investigate their corresponding complements and Lagrangian dual relaxations. We also develop the scalable randomized sampling and local search algorithms with provable performance guarantees. Leveraging the submodularity and concavity of the objective functions in the two proposed formulations, we design a branch and cut algorithm, which can solve the DDF problem to optimality. We further derive a family of optimality cuts, which can significantly enhance the branch and cut performance. Finally, we test our algorithms using real-world data on the new phasor measurement units placement problem in modern power grids considering the existing conventional sensors. Our numerical study demonstrates the efficiency of branch and cut algorithm with optimality cuts and scalability and high-quality outputs of our approximation algorithms.

12 Xinhong Liu, University of Notre Dame, U.S.A.

Optimization of Reactive Ink Formulation for Additive Manufacturing of Charged Membranes

Chemically patterned charged membranes with engendered useful characteristics can offer selective transport of electrolytes. Chemical patterning across the membrane surface via a physical inkjet deposition process requires precise control of the reactive-ink formulation, which enables the introduction of charged functionalities to the membrane. However, the “click” reaction for additive manufacturing of charged membranes involves an excess load of harmful components (i.e., copper), which hinders this type of membrane to be commercialized. This study seeks the optimal reactive ink formulations that consider both precise controllability and environmental concern. We first develop a new dynamic mathematical model for the kinetic reaction network. Nonlinear least-squares parameter estimation is performed to infer unknown kinetic model parameters by analyzing data from multiple dynamic experiments simultaneously. Global sensitivity and Fisher information matrix (FIM) analyses are applied to verify the parameter identifiability and infer additional measurements. Building off the knowledge of the kinetic network, we are able to optimize the reactive ink formulation that maximizes the conversion rate while minimizing the amount of copper. In this poster, we highlight the continuous nonlinear dynamic modeling and computational optimization challenges associated with parameter estimation, design of experiments, and ink formulation.

13 Si Yi Meng, Cornell University, U.S.A.

Reusing function evaluations in derivative-free line search methods

Gradient approximation constructed from function values is crucial in many derivative-free optimization (DFO) algorithms. Finite differencing, Gaussian smoothing, and linear interpolation are among the most popular gradient approximation techniques with extensive studies. In a standard trust-region or line search framework, these methods typically require n function evaluations to compute a new gradient approximation, where n is the problem dimension. When n is large and function evaluations are very expensive, these approaches are rather impractical. Previous works have considered reusing past function evaluations to build linear or quadratic interpolation models in a trust region DFO algorithm. In this work, we focus on backtracking line search and present failure cases when naively applying linear interpolation gradients with recycled past function evaluations. We then propose and analyze modifications to the traditional backtracking to ensure well-poisedness of the sample set. We empirically demonstrate that our method converges faster than using finite-differencing in terms of the number of function evaluations in both synthetic and simulation optimization problems.

14 Wenlong Mou, University of California, Berkeley, U.S.A.
ROOT-SGD: Sharp Nonasymptotics and Asymptotic Efficiency in a Single Algorithm
We study the problem of solving strongly convex and smooth unconstrained optimization problems using stochastic first-order algorithms. We devise a novel algorithm, referred to as Recursive One-Over-T SGD (ROOT-SGD), based on a easily implementable and recursive averaging of past stochastic gradients. We prove that it simultaneously achieves state-of-the-art performance in both a finite-sample, nonasymptotic sense and an asymptotic sense. On the nonasymptotic side, we prove risk bounds on the last iterate of ROOT-SGD with leading-order terms that match the optimal statistical risk with a unity pre-factor, along with a higher-order term that scales at the sharp rate of $\mathcal{O}(n^{-3/2})$. On the asymptotic side, we show that when a mild, one-point Hessian continity condition is imposed, the rescaled last iterate of (multi-epoch) ROOT-SGD converges asymptotically to a Gaussian limit with the Cramér-Rao optimal asymptotic covariance, for a broad range of stepsize choices.

15 Edward Duc Hien Nguyen, Rice University, U.S.A.
Exact Diffusion with Local Steps
This work focuses on solving the decentralized optimization problem in which the objective is to minimize the average of local, private, and non-convex objective functions held by n agents. Communication between agents is governed by a given network topology and agents are restricted in the information they can communicate. Agents can only communicate information such as their parameters and cannot send information such as their local data. State-of-the-art decentralized optimization algorithms such as Decentralized Gradient Descent (D-GD), Exact Diffusion, and Gradient Tracking (GT) developed to solve this problem involve communication between agents of the graph at every iteration. In practice, communication is expensive, resource intensive, and slow. Our goal is to address the issue of communication costs by introducing the technique of local steps into Exact Diffusion. The idea of local steps is to modify these existing algorithms such that the base recursion does not involve any communication between agents for some number of iterations. When a "communication round" occurs, the original recursion with communication is then performed. We analyze the performance of Exact Diffusion with Local Steps analytically and in simulation.

16 Vincent Roulet, University of Washington, U.S.A.
Complexity Bounds of Iterative Linear Quadratic Optimization Algorithms for Discrete Time Nonlinear Control
A classical approach for solving discrete time nonlinear control on a finite horizon consists in repeatedly minimizing linear quadratic approximations of the original problem around current candidate solutions. While widely popular in many domains, such an approach has mainly been analyzed locally. We observe that global convergence guarantees can be ensured provided that the linearized discrete time dynamics are surjective and costs on the state variables are strongly convex. We present how the surjectivity of the linearized dynamics can be ensured by appropriate discretization schemes given the existence of a feedback linearization scheme. We present complexity bounds of algorithms based on linear quadratic approximations through the lens of generalized Gauss-Newton methods. Our analysis uncovers several convergence phases for regularized generalized Gauss-Newton algorithms.

17 Pouya Sampourmahani, Lehigh University, U.S.A.
On the Semidefinite Representation of Second-order Conic Optimization Problems
Second-order conic optimization (SOCO) can be considered as a special case of semidefinite optimization (SDO). Current literature states that a SOCO problem can be embedded in a semidefinite optimization problem using the arrow-head matrix transformation. However, a primal-dual solution pair cannot be mapped simultaneously using the arrow-head transformation as we might lose complementarity and feasibility in some cases. To address this issue, we investigate the relationship between the SOCO problem and its SDO counterpart. Through derivation of standard semidefinite representations of SOCO problems, we introduce an admissible rank-one mapping. We show that the proposed mapping preserves both feasibility and optimality. Next, we discuss how the optimal partition of SOCO problem maps to the optimal partition of

its SDO counterpart. Numerical experiments follow the theoretical proofs for further illustration.

18 Igor Sokolov, King Abdullah University of Science and Technology, Saudi Arabia

EF21: A New, Simpler, Theoretically Better, and Practically Faster Error Feedback

Error feedback (EF), also known as error compensation, is an immensely popular convergence stabilization mechanism in the context of distributed training of supervised machine learning models enhanced by the use of contractive communication compression mechanisms, such as Top-k. First proposed by Seide et al (2014) as a heuristic, EF resisted any theoretical understanding until recently [Stich et al., 2018, Alistarh et al., 2018]. However, all existing analyses either *i)* apply to the single node setting only, *ii)* rely on very strong and often unreasonable assumptions, such global boundedness of the gradients, or iterate-dependent assumptions that cannot be checked a-priori and may not hold in practice, or *iii)* circumvent these issues via the introduction of additional unbiased compressors, which increase the communication cost. In this work we fix all these deficiencies by proposing and analyzing a new EF mechanism, which we call EF21, which consistently and substantially outperforms EF in practice. Our theoretical analysis relies on standard assumptions only, works in the distributed heterogeneous data setting, and leads to better and more meaningful rates. In particular, we prove that EF21 enjoys a fast $\mathcal{O}(1/T)$ convergence rate for smooth nonconvex problems, beating the previous bound of $\mathcal{O}(1/T^{2/3})$, which was shown a bounded gradients assumption. We further improve this to a fast linear rate for PL functions, which is the first linear convergence result for an EF-type method not relying on unbiased compressors. Since EF has a large number of applications where it reigns supreme, we believe that our 2021 variant, EF21, can a large impact on the practice of communication efficient distributed learning. The poster is based on the NeurIPS 2021 proceeding having the same title.

19 Trang Tran, Cornell University, U.S.A.

Policy Optimization for Queueing Models

Queueing systems appear in many important real-life applications including communication networks, transportation and manufacturing systems. In the technology era, these systems have been more and more complex when their underlying dynamics are usually unknown. Reinforcement learning framework is a suitable model for these problems where the agent receives little information from the environment to navigate. In this work, we consider the model-free framework where there is no information on transition probability and the reward function associated with the Markov decision process. We propose a new parameterization of the policy followed by a new algorithm using the intrinsic properties of queueing network systems. We test our method with other standard algorithms.

20 Jie Wang, Georgia Institute of Technology, U.S.A.

Sinkhorn Distributionally Robust Optimization

We study distributionally robust optimization with Sinkhorn distance – a variant of Wasserstein distance based on entropic regularization. We derive convex programming dual reformulations when the nominal distribution is an empirical distribution and a general distribution, respectively. Compared with Wasserstein DRO, it is computationally tractable for a larger class of loss functions, and its worst-case distribution is more reasonable. To solve the dual reformulation, we propose an efficient batch gradient descent with a bisection search algorithm. Finally, we provide various numerical examples using both synthetic and real data to demonstrate its competitive performance.

21 Ke Wang, University of Notre Dame, U.S.A.

Bayesian Optimization Considering Experimental and Physical Constraints – Case Study of Flash Sintering

Solid-state thermoelectric materials convert heat to electricity, having great potential to improve energy efficiency and reduce emission in modern industry. However, the performance of literature reported thermoelectric materials remain inadequate for many industrial applications. Flash sintering, a manufacturing process to produce AgSe based thermoelectric material, heavily relies on intuition driven Edisonian search which is extremely time-consuming. Recent years, there is rising interest in adopting Bayesian optimization (BO) to efficiently optimize the search region of the manufacturing process. BO is a derivative-free optimization framework consisting mainly of two steps: first, a probabilistic surrogate model to emulate the behavior of objective function, e.g., Gaussian Process Regression (GPR); second, an acquisition function, e.g., Expected Improvement (EI), to sequentially recommend future experiments that optimally balance exploitation and exploration. The BO considers five controllable experimental conditions – voltage, pulse duration, number of pulses, pulse delay time and thickness – and one calculated feature, energy, as input features. There are two constraints in the manufacturing process, experiment device constraint and material physics constraint. The first constraint is proposed based on domain knowledge on experiment setup, that there is a certain relation between voltage, pulse duration, and pulse delay time. The second one mainly refers to the boiling point of the AgSe material, and is determined by the energy and thickness, which is proved by simulations in COMSOL. In this poster, we highlight the domain knowledge that can help BO make reasonable recommendations. We also demonstrate how to incorporate machine learning models, e.g., linear regression (LR) and support vector machine (SVM), as constraints to predict if experimental conditions are actually feasible. Finally, we discuss the challenges and opportunities for applying BO to optimize broader manufacturing process.

22 Qi Wang, Lehigh University, U.S.A.
Worst-Case Complexity of TRACE with Inexact Subproblem Solutions for Nonconvex Smooth Optimization
An algorithm for solving nonconvex smooth optimization problems is proposed, analyzed, and tested. The algorithm is an extension of the Trust Region Algorithm with Contractions and Expansions (TRACE) [Math. Prog. 162(1):132, 2017]. In particular, the extension allows the algorithm to use inexact solutions to the arising subproblems, which is an important feature for solving large-scale problems. Inexactness is allowed in a manner such that the optimal iteration complexity of $\mathcal{O}(\epsilon^{-3/2})$ for attaining an ϵ-approximate first-order stationary point is maintained while the worst-case complexity in terms of Hessian-vector products may be significantly improved as compared to the original TRACE. Numerical experiments show the benefits of allowing inexact subproblem solutions and that the algorithm is competitive with state-of-the-art techniques.

23 Zeguan Wu, Lehigh University, U.S.A.
Preconditioned Inexact Infeasible Quantum Interior Point Method for Linear Optimization
Quantum Interior Point Methods (QIPMs) has recently attracted many researchers' interest due to their potential to solve optimization problems significantly faster. Generally, QIPMs, instead of classical linear system solvers use Quantum Linear System Algorithms (QLSAs) to solve the Newton Systems at each iterations. However, the performance of QLSA solvers depends on the condition number of the linear systems, which for QIPMs, typically grows proportional to $\mathcal{O}(1/\mu^2)$, where μ is the central path parameter. To deal with ill-conditioning of the Newton System, a preconditioned QIPM based on the optimal partition estimation is developed. With the preconditioning method, we improve the condition number of the linear systems from $\mathcal{O}(1/\mu^2)$ to $\mathcal{O}(1/\mu)$ and improve the preconditioned QIPM's dependence on complementarity gap from $\mathcal{O}(1/\mu^2)$ to $\mathcal{O}(1/\mu^{1.5})$.

24 Miaolan Xie, Cornell University, U.S.A.
High Probability Iteration and Sample Complexity Bounds for Stochastic Adaptive Step Search
We consider an adaptive step search method for continuous optimization under a stochastic setting where the function values and gradients are available only through inexact probabilistic zeroth and first-order oracles. Unlike the stochastic gradient method and its many variants, the algorithm does not use a pre-specified sequence of step sizes but increases or decreases the step size adaptively according to the estimated progress of the algorithm. The probabilistic oracles capture multiple standard settings including expected loss minimization and zeroth-order optimization. Moreover, our framework allows the function and gradient estimates to be biased. The proposed algorithm is simple to describe and easy to implement. Under reasonable conditions on the oracles, we derive high probability tail bounds on the iteration and sample complexity of the algorithm.

25 Jinwen Yang, University of Chicago, U.S.A.
Nearly Optimal Linear Convergence of Stochastic Primal-Dual Methods for Linear Programming
There is a recent interest on first-order methods for linear programming (LP). In this paper, we propose a stochastic algorithm using variance reduction and restarts for solving sharp primal-dual problems such as LP. We show that the proposed stochastic method exhibits a linear convergence rate for solving sharp instances with a high probability. In addition, we propose an efficient coordinate-based stochastic oracle for unconstrained bilinear problems, which has $\mathcal{O}(1)$ per iteration cost and improves the complexity of the existing deterministic and stochastic algorithms. Finally, we show that the obtained linear convergence rate is nearly optimal (upto log terms) for a wide class of stochastic primal dual methods.

26 Chennan Zhou, Ohio State University, U.S.A.
Effective scenarios in Two-stage DRO: properties and acceleration of decomposition algorithms
Effective scenarios are critical to Distributionally Robust Optimization (DRO) in the sense that, if removed, the optimal value will be changed. Earlier, effective scenarios were examined for convex DRO with Total Variation distance. In this work, we study them for a class of DRO with Wasserstein distance and phi-divergences. We present several properties and discuss how they can be used to accelerate Benders' decomposition algorithms to solve this class of problems.