# Master's Level Internship: Diffusion for Lagrangian Relaxation in MILPs

**Advisors:** Joseph Le Roux (leroux@lipn.fr) and Mathieu Lacroix (lacroix@lipn.fr)
**Location**: LIPN, Université Sorbonne Paris Nord – CNRS
**Duration:** 6 months

## Context

Mixed Integer Linear Programs (MILPs) (Wolsey, 2021) have two main strengths that make them ubiquitous in combinatorial optimization (Korte & Vygen, 2012). First, they can model many combinatorial optimization problems. Second, extremely efficient solvers can now handle MILPs with millions of constraints and variables. They therefore have a wide variety of applications in logistics, telecommunications and beyond. MILP algorithms are exact: they return an optimal solution, or an optimality gap between the returned solution and an optimal one.

MILPs are sometimes hard to solve due to a collection of difficult constraints. Typically, a small number of constraints link together otherwise independent subproblems. For instance, in vehicle routing problems (Golden et al., 2008), there is one independent problem for each vehicle, except for the linking constraints that ensure that exactly one vehicle operates each task of interest. Lagrangian relaxation approaches (Beasley, 1993; Lemaréchal, 2001) are popular in such settings as they allow to unlink the different subproblems. Unfortunately, this relaxation is often solved by iterative methods that require thousands of iterations to recalibrate costs by computing penalties known as Lagrange Multipliers (LMs), to penalise constraint violations while giving a tight bound of the original constrained problem.

Recently (Demelas et al., 2024) introduced a deep learning approach that bypasses the descent, effectively amortizing per instance optimization. A probabilistic encoder based on a graph neural network computes, given a MILP instance and its Continuous Relaxation (CR) solution, high-dimensional representations of relaxed constraints, which are turned into LMs by a decoder. The training parametrises the encoder and the decoder jointly by directly optimizing the bound obtained from the predicted multipliers. Their method is applicable to any problem with a compact MILP formulation, and to any Lagrangian Relaxation providing a tighter bound than CR. Experiments, on two standard problems from the MILP literature, showed that the approach closes up to 85% of the

gap between the continuous relaxation and the best Lagrangian bound, and provides a high-quality warm-start for descent-based Lagrangian methods.

Still, this method is usually unable to cover the whole gap. For this internship, we propose to embed the method defined by (Demelas et al., 2024) into the diffusion framework (Ho et al., 2020), in an unsupervised fashion (Sanokowski et al., 2024).

## Research Proposal

The model of (Demelas et al., 2024) takes as input a representation of the instance to be solved, and candidates Lagrangian multipliers given by the solution of the continuous relaxation of the original problem, and outputs the optimal Lagrangian multipliers (LMs). In other words, we can see this framework as a *denoiser* that, for each instance, returns clean LMs from noisy ones.

From this observation, it is not too hard to derive a diffusion model. Besides the denoiser, or diffusion backward process, that has been discussed above, one has to define the diffusion forward process, that given clean LMs corrupts them to obtain LMs distributed around the CR solution. In this setting, training can be performed following various methods ranging from plain diffusion, to score-based functions (Song et al., 2021), to flow matching (Lipman et al., 2024). However, these settings must be adapted to the unsupervised case: pptimal LMs are not unique and are difficult to collect, hence our choice of unsupervised learning, with a distant learning signal (which can be seen as a reward) given by the relaxed cost of the prediction, see (Sanokowski et al., 2024).

The internship consists of 2 parts :

1. Define the training loss and the inference steps to embed the method of (Demelas et al., 2024) in a diffusion model;
2. Define a neural architecture to implement the diffusion process. We plan to follow the Diffusion Transformer of (Peebles & Xie, 2023), and adapt it to the graph structure of MILPs (Cappart et al., 2021)

## Application

Candidates are expected to be graduating at Master level with a strong background in Machine Learning. Candidates are not required to be knowledgeable in Combinatorial Optimization but a background in this topic will be appreciated.

This is a 6 month internship, to apply please send an email with a CV, a cover letter and a transcript stating clearly your period of availability.

## References

Beasley, J. E. (1993). Lagrangian relaxation. In *Modern heuristic techniques for combinatorial problems* (pp. 243–303). John Wiley & Sons, Inc.

Cappart, Q., Chételat, D., Khalil, E. B., Lodi, A., Morris, C., & Velickovic, P. (2021). Combinatorial optimization and reasoning with graph neural networks. *CoRR*.

Demelas, F., Le Roux, J., Lacroix, M., & Parmentier, A. (2024). Predicting lagrangian multipliers for mixed integer linear programs. *Forty-first International Conference on Machine Learning*. https://openreview.net/forum?id=aZnZOqUOHq

Golden, B., Raghavan, S., & Wasil, E. (Eds.). (2008). *The Vehicle Routing Problem: Latest Advances and New Challenges* (Vol. 43). Springer US.

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *arXiv preprint arxiv:2006.11239*.

Korte, B. H., & Vygen, J. (2012). *Combinatorial optimization: Theory and algorithms* (5th ed). Springer.

Lemaréchal, C. (2001). Lagrangian relaxation. In *Computational combinatorial optimization* (pp. 112–156). Springer.

Lipman, Y., Havasi, M., Holderrieth, P., Shaul, N., Le, M., Karrer, B., Chen, R. T. Q., Lopez-Paz, D., Ben-Hamu, H., & Gat, I. (2024). Flow matching guide and code. https://arxiv.org/abs/2412.06264

Peebles, W., & Xie, S. (2023). Scalable diffusion models with transformers. *Proceedings of the IEEE/CVF international conference on computer vision*, 4195–4205.

Sanokowski, S., Hochreiter, S., & Lehner, S. (2024, July). A diffusion model framework for unsupervised neural combinatorial optimization. In R. Salakhutdinov, Z. Kolter, K. Heller, A. Weller, N. Oliver, J. Scarlett, & F. Berkenkamp (Eds.), *Proceedings of the 41st international conference on machine learning* (pp. 43346–43367, Vol. 235). PMLR. https://proceedings.mlr.press/v235/sanokowski24a.html

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., & Poole, B. (2021). Score-based generative modeling through stochastic differential equations. *International Conference on Learning Representations*. https://openreview.net/forum?id=PxTIG12RRHS

Wolsey, L. A. (2021). *Integer programming* (Second edition). Wiley.