

Post-doc offer

ANR Project **ATHENA**
Sorbonne Université — LPSM

French version beneath.

General Information

Employer	Sorbonne University
Laboratory	LPSM (Laboratory of Probability, Statistics and Modelling)
Supervisors	Sylvain Le Corff, Nicolas Bousquet
Location	Paris, France
Contract	Fixed-term postdoctoral position, 18 to 20 months
Preferred starting date	Q3–Q4 2026
Salary	Approximately €2,900 gross per month, plus reimbursement of half of daily commuting expenses and a housing allowance of approximately €130

Project Overview

The ANR-funded ATHENA project aims to develop methodological and software tools for the automated design of innovative, efficient and low-environmental-impact thermodynamic cycles. It combines superstructure optimisation, generative models and advanced statistical methods to explore complex design spaces for energy processes.

The consortium brings together academic and industrial partners: LRGP at the University of Lorraine, LPSM at Sorbonne University, EDF R&D and Fives ProSim. The overall objective is to develop an open platform for generating, evaluating and optimising robust, safe and high-performance energy systems.

What Is Automated Thermodynamic Cycle Design?

Automated thermodynamic cycle design consists in using algorithms to automatically generate, compare and optimise energy-process architectures, without being restricted to conventional configurations devised by human experts.

A thermodynamic cycle is a sequence of operations—compression, expansion, heat exchange, mixing, separation and others—used to convert energy into electricity, cooling, heating or combinations of these outputs. Traditionally, engineers start from established configurations, such as Brayton cycles, Rankine cycles or heat pumps, and then manually adjust their parameters. Automated design radically broadens this approach.

The process is represented as a **graph**: nodes correspond to components such as turbines, compressors and heat exchangers, while edges represent material and heat flows. Algorithms then explore a very large space of possible graphs to identify structures that are both physically feasible and efficient.

Example 1 — Discovery of non-intuitive configurations A generative algorithm may propose a hybrid architecture combining several thermal loops, for example a Brayton cycle coupled with a waste-heat recovery sub-cycle. Such a configuration may improve energy efficiency by reusing thermal losses that would otherwise be dissipated.

Example 2 — Environmental optimisation Automated design can directly incorporate environmental constraints, including CO₂ emissions, the global warming potential (GWP) of working fluids and resource consumption. The algorithm then searches for cycles that maximise efficiency while minimising environmental impact. For example, it may favour less polluting working fluids or propose architectures that reduce the temperature of discharged waste heat.

Example 3 — Adaptation to a local context At a given industrial site, the available hot and cold sources impose specific constraints. Automated generation can produce cycles tailored to these local conditions by exploiting waste heat or renewable energy sources. Manually identifying such configurations would be difficult within a very large solution space.

In principle, automated design makes it possible to:

- systematically explore millions of possible configurations;
- reveal non-intuitive solutions;
- simultaneously account for energy, economic and environmental objectives;
- accelerate the transition towards more resource-efficient and cleaner energy systems.

This topic has long been the subject of extensive research involving a range of optimisation and exploration techniques; see, for example, [13, 6, 8, 5, 4, 14, 1, 9, 3]. These approaches still face major computational challenges. More recently, advances in embedding methods into continuous latent spaces and in diffusion models have opened up new and promising avenues for overcoming these difficulties.

Postdoctoral Research Topic

The postdoctoral position is part of **Work Package 1** of the ANR project, which focuses on the targeted generation of processes. The scientific objective is to develop *stochastic preconditioning* methods capable of intelligently restricting the structural design space of thermodynamic cycles before performing computationally expensive optimisation.

The processes are therefore represented as graphs. The resulting combinatorial space is extremely large, non-convex and potentially discontinuous. The work will exploit VAE-type autoencoding methods, which provide an embedding into a continuous latent space with an effective dimension of approximately a few dozen variables. The purpose is to facilitate the exploration of this space and the identification of a domain of feasible processes. The expected benefit is twofold: to improve the guidance of optimisation algorithms and to contribute to the development of improved optimisation procedures. Initial explorations of this latent space reveal disconnected regions and a complex feasibility landscape within the available data. This observed landscape is therefore itself constrained by the distribution of the available data.

The work will rely on datasets containing several tens of thousands of training examples, produced in particular through simulations of energy cycles such as Brayton cycles, Rankine cycles and heat pumps. It will also build on a body of literature arising primarily from previous work conducted at EDF and LRGP, including [12, 11, 10], as well as on LPSM expertise in stochastic exploration and deep learning; see, for example, [2, 15]. Two main research directions will be prioritised:

- incorporating constraints into the construction of variational-autoencoder priors by conditioning them on known data; quantisation-based approaches such as VQ-VAE [7], traditionally associated with image processing, may also be investigated;
- developing specific diffusion models and algorithms for exploring new and potentially relevant configurations.

These methodological developments will necessarily involve software implementation and interaction with the software tools developed in the other work packages of the project. Particular attention will be paid to reproducibility, software integration, industrial validation and the dissemination of

the proposed methods within the academic community.

At LPSM, the postdoctoral researcher will benefit from the methodological support of a growing community specialising in machine learning, *Machine Learning at LPSM*, as well as from the dedicated artificial-intelligence environment provided by the *Sorbonne Center for Artificial Intelligence* (SCAI), Sorbonne University's AI institute, of which LPSM is a member. The work will also require close interaction with the other laboratories involved in the ANR ATHENA project: EDF R&D's LNHE laboratory in the Paris region and LRGP at the University of Lorraine. Travel and research stays at these partner institutions should therefore be expected.

Main Responsibilities

- Develop machine-learning methods for exploring process design spaces
- Design embedding models and generative models
- Implement the proposed methods and integrate them into the ATHENA platform
- Perform numerical validation on industrial case studies
- Write scientific publications and disseminate the research results

Candidate Profile

- PhD in applied mathematics, statistics, computer science or a related field
- Strong background in statistical learning, particularly embedding and diffusion methods, and ideally in optimisation
- Proficiency in Python and scientific software engineering
- Interest in energy processes and mathematical modelling
- Ability to work effectively in a collaborative research environment

Application

Applicants should send a CV, a list of publications, references, and links to relevant code or datasets to:

`nicolas.bousquet@sorbonne-universite.fr` and `sylvain.le-corff@sorbonne-universite.fr`

FRENCH VERSION

Informations générales

Employeur	Sorbonne Université
Laboratoire	LPSM (Laboratoire de Probabilité, Statistique et Modélisation)
Encadrants	Sylvain Le Corff, Nicolas Bousquet
Localisation	Paris, France
Contrat	CDD post-doctoral, 18 à 20 mois
Début souhaité	Q3-Q4 2026
Rémunération	~ 2900 € brut + prise en charge de la moitié des frais de transport quotidiens + allocation logement (~ 130 €)

Présentation du projet

Le projet ANR ATHENA vise à développer des outils méthodologiques et logiciels pour la conception automatisée de cycles thermodynamiques innovants, efficaces et à faible impact environnemental. Il combine optimisation superstructurale, modèles génératifs et méthodes statistiques avancées afin d'explorer des espaces de conception complexes pour les procédés énergétiques.

Le consortium regroupe des partenaires académiques et industriels (LRGP à l'Université de Lorraine, LPSM à Sorbonne Université, EDF R&D, Fives ProSim). L'objectif global est la création d'une plateforme ouverte permettant la génération, l'évaluation et l'optimisation de systèmes énergétiques robustes, sûrs et performants.

Qu'est-ce que la conception automatisée de cycles thermodynamiques ?

La conception automatisée de cycles thermodynamiques consiste à utiliser des algorithmes pour générer, comparer et optimiser automatiquement des architectures de procédés énergétiques, sans se limiter aux configurations classiques imaginées par un expert humain.

Un cycle thermodynamique est un enchaînement d'opérations (compression, détente, échange thermique, mélange, séparation, etc.) permettant de transformer de l'énergie : produire de l'électricité, du froid, du chauffage ou des combinaisons de ces usages. Traditionnellement, l'ingénieur part de schémas connus (cycle de Brayton, cycle de Rankine, pompe à chaleur) puis ajuste manuellement les paramètres. La conception automatisée élargit radicalement cette approche.

Le procédé est représenté comme un **graphe** : les nœuds correspondent aux équipements (turbine, compresseur, échangeur) et les arêtes aux flux de matière et de chaleur. Les algorithmes explorent ensuite un espace très large de graphes possibles afin d'identifier des structures physiquement réalisables et performantes.

Exemple 1 — Découverte de configurations non intuitives Un algorithme génératif peut proposer une architecture hybride combinant plusieurs boucles thermiques, par exemple un cycle Brayton couplé à un sous-cycle de récupération de chaleur. Ce type de configuration peut améliorer le rendement énergétique en réutilisant des pertes thermiques qui seraient autrement dissipées.

Exemple 2 — Optimisation environnementale La conception automatisée peut intégrer directement des contraintes environnementales : émissions de CO₂, potentiel de réchauffement global (GWP) des fluides, consommation de ressources. L'algorithme recherche alors des cycles qui maximisent l'efficacité tout en minimisant l'impact environnemental. Par exemple, il peut privilégier des fluides de travail moins polluants ou proposer des architectures réduisant la température des rejets thermiques.

Exemple 3 — Adaptation à un contexte local Dans un site industriel donné, les sources chaudes et froides disponibles imposent des contraintes spécifiques. La génération automatisée peut produire des cycles adaptés à ces conditions locales, en tirant parti de chaleurs fatales ou de sources renouvelables, ce qui serait difficile à concevoir manuellement dans un espace de solutions très vaste.

En résumé, la conception automatisée permet en théorie :

- d’explorer systématiquement des millions de configurations possibles ;
- de révéler des solutions non intuitives ;
- d’intégrer simultanément des objectifs énergétiques, économiques et environnementaux ;
- d’accélérer la transition vers des systèmes énergétiques plus sobres et plus propres.

Ce sujet fait l’objet, depuis longtemps, de nombreuses recherches mobilisant plusieurs techniques d’optimisation et d’exploration différentes (ex : [13, 6, 8, 5, 4, 14, 1, 9, 3]). Elles se heurtent encore à des difficultés computationnelles fortes. Récemment, l’élan donné aux méthodes de plongement vers des espaces latents continus, et aux modèles de diffusion, offrent des perspectives nouvelles et prometteuses pour résoudre ces difficultés.

Sujet du post-doctorat

Le post-doctorat s’inscrit dans le **Work Package 1** du projet ANR, consacré à la génération ciblée de procédés. L’objectif scientifique est le développement de méthodes de *préconditionnement stochastique* permettant de restreindre intelligemment l’espace structurel des cycles thermodynamiques avant optimisation coûteuse.

Les procédés sont donc représentés comme des graphes. L’espace combinatoire est très grand, non convexe et potentiellement discontinu. Le travail consiste à utiliser les bénéfices de méthodes d’autoencodage de type VAE, proposant un plongement dans un espace latent continu (de dimension effective de l’ordre de quelques dizaines), pour faciliter l’exploration et la délimitation d’un domaine de faisabilité des procédés. Le bénéfice attendu est de pouvoir améliorer le guidage d’algorithmes d’optimisation, et de pouvoir également contribuer à améliorer ces algorithmes. Les premières explorations de cet espace latent montrent des pertes de connexité et un paysage complexe du sous-ensemble de faisabilité, au travers des données disponibles (ce paysage étant donc limité par la distribution des données disponibles).

Le travail se fondera justement sur des bases de plusieurs dizaines de milliers de données d’apprentissage, produites notamment par simulation de cycles énergétiques (Brayton, Rankine, pompes à chaleur), et une littérature essentiellement issue de travaux historiques d’EDF et du LRGP (notamment [12, 11, 10]) et de compétences en exploration stochastique et en apprentissage profond portées par le LPSM (ex : [2, 15]). Deux pistes seront privilégiées :

- insérer des contraintes dans la construction du prior des autoencodeurs variationnels par conditionnement sur les données connues ; des approches de quantization du type VQ-VAE [7], que l’on associe traditionnellement au traitement d’image, pourront être investiguées ;
- construire des modèles (algorithmes) de diffusion spécifiques permettant d’explorer de nouvelles configurations possiblement pertinentes.

Ces développements s’appuieront nécessairement sur du développement informatique qui devra interagir avec les sur les outils logiciels développés dans les autres lots du projet. Une attention particulière sera portée à la reproductibilité, à l’intégration logicielle et à la validation industrielle, ainsi qu’à la diffusion des méthodes au sein de la communauté académique.

Le ou la post-doctorant(e) bénéficiera du soutien méthodologique, au LPSM, d’une communauté croissante, spécialisée en machine learning (*Machine Learning at LPSM*) et de l’environnement dédié "IA" apporté par le *Sorbonne Center for Artificial Intelligence* (SCAI), institut d’IA de Sorbonne Université (dont est membre le LPSM). Par ailleurs, ce travail nécessitera des fortes interactions

avec les autres laboratoires du projet ANR ATHENA : le LNHE d'EDF R&D en région parisienne, et le LRGP de l'Université de Lorraine. Des déplacements et des séjours de travail seront donc à prévoir.

Missions principales

- Développement de méthodes d'apprentissage pour l'exploration d'espaces de procédés
- Conception de modèles d'embedding et modèles génératifs
- Implémentation logicielle et intégration dans la plateforme ATHENA
- Validation numérique sur cas industriels
- Rédaction d'articles scientifiques et diffusion des résultats

Profil recherché

- Doctorat en mathématiques appliquées, statistique, informatique ou domaine connexe
- Solide expérience en apprentissage statistique (plongement, méthodes de diffusion) et idéalement en optimisation
- Maîtrise de Python et ingénierie logicielle scientifique
- Intérêt pour les procédés énergétiques et la modélisation
- Capacité à travailler en environnement collaboratif

Candidature

Envoyer CV, liste de publications, références et liens vers code ou données à :

`nicolas.bousquet@sorbonne-universite.fr` et `sylvain.le-corff@sorbonne-universite.fr`

References

- [1] A. Chun, J. L. M. Donatelli, J. J. C. S. Santos, C. B. Zabeu, and M. Carvalho. Superstructure optimization of absorption chillers integrated with a large internal combustion engine for waste heat recovery and repowering applications: Thermodynamic and economic assessments. *Energy*, 263:125970, 2023.
- [2] Max Cohen, Guillaume Quispe, Sylvain Le Corff, Charles Ollion, and Eric Moulines. Diffusion bridges vector quantized variational autoencoders. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 4141–4156. PMLR, 17–23 Jul 2022.
- [3] Qinghe Gao, Haoyu Yang, Maximilian F. Theisen, and Artur M. Schweidtmann. Accelerating process synthesis with reinforcement learning: Transfer learning from multi-fidelity simulations and variational autoencoders. *Computers & Chemical Engineering*, 2025.
- [4] Quirin G"ottl, Yannic T"onges, Dominik G. Grimm, and Jakob Burger. Automated flowsheet synthesis using hierarchical reinforcement learning: Proof of concept. *Chemie Ingenieur Technik*, 93(12):1929–1938, 2021.
- [5] Ahmad Khan and Alexei Lapkin. Searching for optimal process routes: A reinforcement learning approach. *Computers & Chemical Engineering*, 141:107027, 2020.
- [6] Luca Mencarelli, Qi Chen, Alexandre Pagot, and Ignacio E. Grossmann. A review on superstructure optimization approaches in process system engineering. *Computers & Chemical Engineering*, 136:106808, 2020.
- [7] Fabian Mentzer, David Minnen, Eirikur Agustsson, and Michael Tschannen. Finite scalar quantization: Vq-vae made simple. In *International Conference on Learning Representations*, volume 2024, pages 51772–51783, 2024.
- [8] Laurence I. Midgley. Deep reinforcement learning for process synthesis. *arXiv preprint arXiv:2009.13265*, 2020.
- [9] Efstratios N. Pistikopoulos and Yuhe Tian. Advanced modeling and optimization strategies for process synthesis. *Annual Review of Chemical and Biomolecular Engineering*, 15:81–103, 2024.
- [10] Antonio Rocha Azevedo. *Methodology for systematic process design: analysis of superstructural and generative approaches*. PhD thesis, Université de Lorraine, 2025.
- [11] Antonio Rocha Azevedo, Tahar Nabil, Valentin Loubière, Romain Privat, Thibaut Neveux, and Jean-Marc Commenge. Comparing generative process synthesis approaches with superstructure optimization for the conception of supercritical co₂ brayton cycles. *Computers & Chemical Engineering*, 201:109255, 2025.
- [12] Antonio Rocha Azevedo, Tahar Nabil, Valentin Loubière, Benoît Valentin, Romain Privat, Thibaut Neveux, and Jean-Marc Commenge. Representation learning for flowsheets: generating structures for process synthesis. 2024.
- [13] Joonjae Ryu, Lingxun Kong, Arthur E. Pastore de Lima, and Christos T. Maravelias. A generalized superstructure-based framework for process synthesis. *Computers & Chemical Engineering*, 129:106481, 2019.
- [14] Laura Stops, Roel Leenhouts, Qinghe Gao, and Artur M. Schweidtmann. Flowsheet generation through hierarchical reinforcement learning and graph neural networks. *AIChE Journal*, 69(1):e17938, 2023.

- [15] S. Surendran, A. Fermanian, and S. Le Corff. Latent Guided Sampling for Combinatorial Optimization. *arXiv:2506.03672v1*, 2025.