

Titre: Apprentissage profond guidé par la physique pour la modélisation
de la dynamique de points chauds des rubans supraconducteurs

Auteur: Nicolas Breault
Author:

Date: 2025

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Breault, N. (2025). Apprentissage profond guidé par la physique pour la
modélisation de la dynamique de points chauds des rubans supraconducteurs
Citation: [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie.
<https://publications.polymtl.ca/67860/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/67860/>
PolyPublie URL:

**Directeurs de
recherche:** Frédéric Sirois
Advisors:

Programme: Génie électrique
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Apprentissage profond guidé par la physique pour la modélisation de la
dynamique de points chauds des rubans supraconducteurs**

NICOLAS BREault

Département de génie électrique

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Génie électrique

Août 2025

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

**Apprentissage profond guidé par la physique pour la modélisation de la
dynamique de points chauds des rubans supraconducteurs**

présenté par **Nicolas BREault**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

Serge PRUDHOMME, président

Frédéric SIROIS, membre et directeur de recherche

Philippe FAZILLEAU, membre

DÉDICACE

À Léo, mon chat, présent depuis le début de mes études universitaires, mais qui n'a malheureusement pas eu la chance d'en voir l'aboutissement . . .

REMERCIEMENTS

Tout d’abord, j’aimerais sincèrement remercier Frédéric Sirois, mon directeur de recherche, et Christian Lacroix, associé de recherche que je considère informellement comme mon codirecteur, pour ces deux magnifiques années de maîtrise. Un grand merci à vous deux de m’avoir proposé ce projet de recherche qui m’a passionné, alliant modélisation numérique et apprentissage profond, deux sujets d’intérêt que j’ai pu combiner en un seul projet. Vous êtes deux experts exemplaires avec qui il fut un réel plaisir de travailler, et avec qui les discussions ont été des plus intéressantes. Votre enthousiasme face à mes résultats et les discussions ouvertes qui ont suivi ont été une très grande source de motivation pour me dépasser. Vous étiez, sincèrement, un des meilleurs duos pour m’assister dans ce projet. Par ailleurs, j’aimerais une fois de plus remercier Frédéric pour m’avoir donné l’opportunité de voyager en Suisse dans le cadre d’une conférence en lien avec mon projet de recherche. Il s’agissait de mon tout premier voyage depuis 6 ans, mais surtout de ma toute première fois en Europe. Je n’oublierai jamais ce voyage que j’ai adoré, merci énormément pour cette chance.

Je remercie également du fond du coeur le Fonds de recherche du Québec – Nature et technologies (FRQNT) pour leur financement, sans quoi ce projet n’aurait pas été possible.

Dans le même ordre d’idées, je remercie les membres du jury, Serge Prudhomme, dont le cours d’éléments finis fut particulièrement stimulant, ainsi que Philippe Fazilleau, chercheur au CEA à Saclay, France, d’avoir accepté de réviser ce mémoire. J’espère profondément que ce mémoire saura refléter l’intérêt que j’ai eu pour ce projet.

Un grand merci à tous les membres du laboratoire de supraconductivité et magnétisme que j’ai côtoyés au cours de ces deux dernières années. Certes, je n’ai pas été l’étudiant le plus présent au laboratoire, mais sachez qu’il fut toujours un plaisir de partager mon heure de lunch avec vous. De même, j’aimerais particulièrement remercier Gregory Giard, qui a joué un rôle clé dans l’organisation de la conférence en Suisse, et avec qui j’ai voyagé en Autriche et en Allemagne par la suite. Bien évidemment, je tiens à remercier Simon-Mathieu Bergeron Hartmann, un ami du baccalauréat. Nos parties de boule de papier ainsi que nos discussions très pertinentes furent fort appréciées. Nous étions certainement une source de distraction mutuellement bénéfique.

Enfin, j’aimerais remercier tous mes amis proches ainsi que ma famille, dont Christian, Caroline, Ricardo, Raphaëlle, et Denise pour leur soutien inconditionnel tout au long de mes études. J’espère, pour ceux qui assisteront à ma soutenance, que je pourrai finalement répondre à vos interrogations quant au sujet de mes travaux.

Pour terminer, j'aimerais remercier ma copine Éloïse qui a toujours cru en moi, et je dois l'admettre, bien souvent plus que moi-même.

RÉSUMÉ

La modélisation de la dynamique de points chauds dans les rubans supraconducteurs à haute température critique (HTS) requiert des méthodes numériques robustes, en raison de la résolution de systèmes d'équations différentielles hautement non linéaires, de l'usage de maillages irréguliers, et de la présence de systèmes matriciels mal conditionnés. Bien que de telles méthodes numériques existent aujourd'hui, elles ont toutefois l'inconvénient majeur d'être sujettes à des instabilités numériques et exigeantes en temps de calcul, principalement en raison de leur formulation bidimensionnelle (2-D) et tridimensionnelle (3-D). Comme la modélisation du quench dans les rubans supraconducteurs requiert de telles formulations afin de reproduire la physique observée expérimentalement, il est alors impraticable de simuler des rubans HTS aux dimensions réelles.

Afin de lever en partie ces erreurs, ce mémoire a comme objectif de concevoir un modèle numérique réduisant significativement les coûts de calcul des méthodes numériques traditionnelles afin de simuler la dynamique du quench de rubans HTS de grande longueur. Pour ce faire, le modèle proposé repose d'une part sur une réduction de dimensionnalité, formulant alors la dynamique 2-D/3-D des rubans supraconducteurs en un problème électrothermique 1-D, et d'autre part, le modèle fait intervenir des méthodes d'apprentissage profond afin d'apprendre à compenser l'erreur associée à cette réduction de dimensionnalité. Le modèle hybride proposé permet donc d'apprendre la solution électrothermique longitudinale du ruban à un coût réduit sans pour autant sacrifier la précision.

Ceci dit, le modèle d'apprentissage profond guidé par la physique développé dans ce travail suppose qu'on s'intéresse uniquement à la dynamique longitudinale au centre des couches supraconductrice ou stabilisatrice. Cette hypothèse est justifiée en raison des dimensions particulières des rubans HTS, s'apparentant à une géométrie quasi-planaire où l'épaisseur est de l'ordre des micromètres, la largeur est de l'ordre des millimètres, et la longueur est de l'ordre des mètres, voire des centaines de mètres.

Une fois le modèle numérique introduit en détail, ce dernier est entraîné à partir d'un jeu de données de ruban HTS uniforme formé d'une centaine de simulations par éléments finis 2-D. Le modèle numérique 1-D est alors capable de reproduire avec précision la solution éléments finis au centre de la couche HTS en température, en densité de courant ainsi qu'en potentiel électrique en seulement 12 secondes, soit deux ordres de grandeur de moins que les éléments finis (facteur de 450). Le modèle de référence, basé sur la méthode des différences finies 1-D, est quant à lui imprécis et peu fiable, notamment en raison des effets multidi-

mensionnels qui lui échappent. Le modèle proposé se démarque également par sa capacité à étendre son domaine, lui permettant alors d’allonger ou rétrécir la longueur du ruban une fois entraîné. Cette propriété particulièrement intéressante permet alors « d’apprendre » la physique sur une échelle de longueur de quelques centimètres, et par la suite de prédire la dynamique du quench sur des rubans de plusieurs mètres. Celle-ci a été mise en évidence en entraînant le modèle sur des rubans de longueur de 7 centimètres, puis en prédisant ensuite le comportement d’un ruban de plus d’un 1.4 mètre avec exactitude en seulement 1.6 minute.

Dans le but de pousser le projet encore plus loin, le modèle numérique est entraîné sur des données de ruban CFD issues de dizaines simulations 3-D par éléments finis. Les rubans CFD, développés à Polytechnique Montréal et récompensés en 2025 par le prix Honoris Genius – Innovation technologique de l’ordre des ingénieurs du Québec, ont la particularité d’être des rubans HTS engendrant des effets 3-D extrêmement importants. Au meilleur de notre connaissance, aucun modèle 1-D classique ne peut répliquer leur comportement. Toutefois, il a été démontré qu’après certaines modifications ingénieuses du modèle numérique proposé, celui-ci peut reproduire la physique 3-D inhérente des rubans CFD à partir d’une formulation unidimensionnelle hybride. Le modèle proposé présente ainsi des performances remarquables quant aux solutions longitudinales en température, en densité de courant, et en potentiel électrique, tout en réduisant le temps de calcul de plus de 3 ordres de grandeur (facteur de 4000).

Il ressort donc de ce mémoire le tout premier modèle 1-D en mesure de reproduire fidèlement la dynamique du quench dans les rubans uniformes 2-D et les rubans CFD 3-D. Un tel modèle permet alors de simuler le comportement des rubans HTS aux propriétés microscopiques dans des logiciels de réseaux électriques pour des applications à grande échelle, ou même de jouer le rôle d’un outil robuste pour le prototypage de rubans, de câbles ou même d’aimants supraconducteurs. Par ailleurs, le modèle proposé permettrait dans un futur proche de modéliser efficacement, et pour la toute première fois, des bobines supraconductrices à base de ruban CFD.

ABSTRACT

Modeling the dynamics of hot spots in high-temperature superconducting (HTS) tapes requires robust numerical methods due to the resolution of highly nonlinear differential equation systems, the use of irregular meshes, and the presence of ill-conditioned matrix systems. Although such numerical methods are currently available, they suffer from a major drawback: they are prone to numerical instabilities and computationally demanding, primarily due to their two-dimensional (2-D) and three-dimensional (3-D) formulations. Since modeling quench phenomena in superconducting tapes requires such formulations to accurately reproduce experimentally observed physics, simulating HTS tapes at real-world scales becomes impractical.

To partially overcome these limitations, this thesis aims to design a numerical model that significantly reduces the computational cost of traditional numerical methods, in order to simulate the quench dynamics of long-length HTS tapes. To achieve this, the proposed model relies on a dimensionality reduction, formulating the 2-D/3-D dynamics of superconducting tapes as a 1-D electrothermal problem, and on the other hand, the model uses deep learning methods to learn how to compensate for the error associated with this dimensionality reduction. The proposed hybrid model thus makes it possible to learn the tape's longitudinal electrothermal solution at reduced cost without sacrificing accuracy.

That said, the physics-guided deep learning model developed in this work assumes that we are only interested in the longitudinal dynamics at the center of the superconducting or stabilizing layers. This assumption is justified by the particular dimensions of HTS tapes, resembling a quasi-planar geometry where thickness is of the order of micrometers, width is of the order of millimeters, and length is of the order of meters or even hundreds of meters.

Once the numerical model is introduced in detail, it is trained using a dataset of uniform HTS tapes composed of about one hundred 2-D finite element simulations. The resulting 1-D numerical model is then able to accurately reproduce the finite element solution at the center of the HTS layer in terms of temperature, current density, and electric potential in only 12 seconds—two orders of magnitude faster than the finite element method (a speed-up factor of 450). In contrast, the reference model based on the 1-D finite difference method is inaccurate and unreliable, particularly due to the multidimensional effects it fails to capture. That said, the proposed model also stands out for its ability to extend its spatial domain, allowing it to simulate longer tape lengths once trained. This particularly valuable feature enables the model to "learn" the underlying physics on a spatial scale of just a few centimeters, and

then accurately predict quench dynamics over tapes several meters long. This capability was demonstrated by training the model on 7-centimeter tapes and subsequently predicting the behavior of a 1.4-meter-long tape with high accuracy in just 1.6 minutes.

To take the project a step further, the numerical model is trained on CFD tape data generated from dozens of 3-D finite element simulations. Developed at Polytechnique Montréal and awarded the Honoris Genius – Innovation technologique by the Ordre des ingénieurs du Québec in 2025, CFD tapes are a class of HTS conductors that exhibit pronounced 3-D effects—so significant that, to the best of current knowledge, no 1-D model has been able to replicate their behavior. However, it is demonstrated that, following a set of carefully designed modifications, the proposed numerical model can effectively capture the intrinsic 3-D physics of CFD tapes using a one-dimensional hybrid formulation. The model thus achieves remarkable accuracy in predicting the longitudinal distributions of temperature, current density, and electric potential, while reducing computation time by more than three orders of magnitude (a factor of 4000).

This thesis thus presents the very first 1-D model capable of faithfully reproducing quench dynamics in both 2-D uniform tapes and 3-D CFD tapes. Such a model enables the simulation of HTS tape behavior with microscale properties within power system software for large-scale applications, or even serves as a robust tool for the prototyping of superconducting tapes, cables, or magnets. Furthermore, the proposed model could, in the near future, make it possible—for the very first time—to efficiently simulate superconducting coils based on CFD tapes.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	vi
ABSTRACT	viii
TABLE DES MATIÈRES	x
LISTE DES TABLEAUX	xiv
LISTE DES FIGURES	xv
LISTE DES SIGLES ET ABRÉVIATIONS	xvii
LISTE DES ANNEXES	xxi
CHAPITRE 1 INTRODUCTION	1
1.1 Mise en contexte	1
1.2 Problématique	1
1.3 Objectifs de recherche	2
1.4 Plan du mémoire	3
1.5 Contributions de ce mémoire	3
CHAPITRE 2 LES RUBANS SUPRACONDUCTEURS	4
2.1 La supraconductivité : définitions et propriétés	4
2.1.1 Les paramètres critiques	4
2.1.2 Surface critique	4
2.1.3 Les LTS et HTS	5
2.1.4 Les supraconducteurs de type I et II	5
2.1.5 La relation $E - J$ des HTS	6
2.2 Les caractéristiques des rubans supraconducteurs	7
2.2.1 La vitesse de propagation de la zone normale (NZPV)	7
2.2.2 La longueur de transfert du courant (CTL)	9
2.2.3 La résistance d'interface	10

2.2.4	Stabilité thermique	10
2.3	Les architectures des rubans supraconducteurs	11
2.3.1	Architecture uniforme	11
2.3.2	Architecture CFD	12
2.4	La modélisation des rubans 2G HTS CC	13
2.4.1	Modèle 1-D pour la dynamique de quench	14
2.4.2	Limitations du modèle analytique 1-D	16
CHAPITRE 3 TECHNIQUES DE MODÉLISATION : DE L'APPRENTISSAGE PRO-		
	FOND À LA PHYSIQUE	17
3.1	Notions de base en apprentissage profond	17
3.1.1	Les réseaux de neurones	17
3.1.2	Les réseaux récurrents	18
3.1.3	Le mécanisme d'attention	19
3.1.4	L'entraînement de modèles	20
3.1.5	Le surapprentissage	22
3.2	L'apprentissage profond guidé par la physique	22
3.2.1	Les modèles hybrides	23
3.2.2	Les réseaux de neurones informés par la physique (PINN)	24
3.2.3	Les opérateurs neuronaux profonds (DeepONet)	25
3.3	Modèle hybride appliqué aux problèmes thermiques	25
3.3.1	La discrétisation de Du Fort-Frankel	26
3.3.2	Le modèle des coefficients effectifs (MCE)	27
3.4	La supraconductivité par apprentissage profond	28
CHAPITRE 4 MODÉLISATION DE LA DYNAMIQUE DE QUENCH PAR L'AP-		
	PRENTISSAGE PROFOND	31
4.1	Stratégies d'accélération des calculs	31
4.1.1	Remplacer les méthodes traditionnelles par les méthodes d'apprentis- sage profond	31
4.1.2	Réduire la dimensionnalité du problème	32
4.1.3	La stratégie retenue	32
4.2	Modélisation de la dynamique de quench par éléments finis	33
4.2.1	Simulations par éléments finis	34
4.2.2	Génération des données d'entraînement	37
4.3	Modèle de référence	42

4.4	Modèle d'apprentissage profond guidé par la physique pour la dynamique de quench	45
4.4.1	Inconvénients du modèle des coefficients effectifs	46
4.4.2	Reformulation du modèle des coefficients effectifs	47
4.4.3	Conservation de la chaleur	48
4.4.4	Le modèle RAPDi	49
4.4.5	Propriété d'extension du domaine	51
4.4.6	Stratégies d'entraînement	53
CHAPITRE 5 MODÉLISATION DU QUENCH DANS DES RUBANS D'ARCHITECTURE 2-D UNIFORME		55
5.1	Optimisation de l'architecture et des performances du modèle RAPDi	55
5.1.1	L'espace de modèles RAPDi étudié	56
5.1.2	Performances selon les métriques	58
5.1.3	Performances temporelles	61
5.2	Solutions en température d'un ruban uniforme	61
5.2.1	Solution thermique pour le régime dominé par les pertes Joule	62
5.2.2	Solution thermique pour le régime dominé par le refroidissement	63
5.3	Solutions électriques d'un ruban uniforme	66
5.3.1	Solution en densité de courant	66
5.3.2	Solution en potentiel électrique	67
5.4	Modélisation d'un ruban HTS uniforme de grande longueur	69
CHAPITRE 6 MODÉLISATION DE RUBANS CFD D'ARCHITECTURE 3-D		72
6.1	Du ruban CFD 3-D au ruban 1-D : les considérations	72
6.1.1	Distribution de température	73
6.1.2	Résistance d'interface effective	75
6.1.3	Longueur de transfert	76
6.2	Génération de données	78
6.2.1	Paramètres géométriques et inhomogénéités aléatoires	78
6.2.2	Rampes de courant aléatoires	79
6.3	Modifications des modèles 1-D	81
6.3.1	Modèle de référence	81
6.3.2	Modèle RAPDi	82
6.4	Architecture et performances du modèle RAPDi	83
6.4.1	Espace de modèles étudiés	84
6.4.2	Performances des modèles étudiés	85

6.5	Solutions électrothermiques d'un ruban CFD	88
6.5.1	Solution du problème thermique	88
6.5.2	Solution du problème électrique	90
CHAPITRE 7 CONCLUSION		94
7.1	Synthèse des travaux	94
7.2	Limitations de la solution proposée	96
7.3	Améliorations futures	96
RÉFÉRENCES		98
ANNEXES		103

LISTE DES TABLEAUX

Tableau 4.1	Paramètres géométriques de rubans HTS uniformes.	38
Tableau 4.2	Paramètres du REBCO.	39
Tableau 5.1	Modèles RAPDi étudiés à base de réseau <i>feed forward</i>	56
Tableau 5.2	Modèles RAPDi étudiés à base d'apprentissage profond avec attention.	56
Tableau 5.3	Erreurs des modèles RAPDi étudiés à base de réseau <i>feed forward</i> . . .	59
Tableau 5.4	Erreurs des modèles RAPDi étudiés à base d'APA.	59
Tableau 5.5	Temps de calcul moyen d'une simulation et le facteur d'accélération obtenu par rapport à la MEF 2-D pour différents modèles.	61
Tableau 5.6	Erreur relative maximale de la solution $V_s(x, t)$ pour un ruban uniforme à différents instants pour la MDF 1-D et le modèle RAPDi.	69
Tableau 5.7	Temps de calcul du modèle RAPDi 1-D pour différentes longueurs de ruban.	71
Tableau 6.1	Paramètres géométriques de rubans CFD enrobés.	79
Tableau 6.2	Modèles RAPDi étudiés à base de réseau <i>feed forward</i>	84
Tableau 6.3	Modèles RAPDi étudiés à base de modèles d'APA.	84
Tableau 6.4	Erreurs des modèles RAPDi étudiés à base de réseau <i>feed forward</i> . . .	86
Tableau 6.5	Erreurs des modèles RAPDi étudiés à base d'APA.	86
Tableau 6.6	Temps de calcul moyen d'une simulation et facteur d'accélération obtenu par rapport à la MEF 3-D pour différents modèles.	87
Tableau 6.7	Erreur relative maximale de la solution $V_s(x, t)$ pour un ruban CFD .	93
Tableau C.1	Erreur relative maximale de la solution $V_s(x, t)$ pour un ruban CFD .	107

LISTE DES FIGURES

Figure 2.1	Surface critique d'un supraconducteur dans l'espace (T, B, J)	5
Figure 2.2	Comparaison des supraconducteurs de type I et type II	6
Figure 2.3	Relation non linéaire $E - J$ d'un supraconducteur de type II	7
Figure 2.4	Distribution réelle de courant critique I_c	8
Figure 2.5	Vue longitudinale d'un partage de courant à l'intérieur d'un ruban HTS	9
Figure 2.6	NZPV en fonction du courant I pour les 3 régimes de stabilité thermique	11
Figure 2.7	Vue transverse d'un ruban 2G HTS CC uniforme et CFD	12
Figure 2.8	Comparaison entre l'architecture uniforme et l'architecture CFD . . .	13
Figure 2.9	Exemple d'un ruban supraconducteur uniforme 3-D en W/m^3 et le ruban équivalent en W/m^2 selon le modèle de Levin <i>et al.</i>	16
Figure 3.1	Schéma d'un réseau de neurones	18
Figure 3.2	Schéma d'un réseau de neurones récurrent	19
Figure 3.3	Schéma d'un RNN récurrent entraîné à l'aide du forçage par enseignant	21
Figure 3.4	Schéma d'un modèle d'AP formé de divises composantes	23
Figure 3.5	Schéma de différents modèles d'apprentissage profond guidé par la phy- sique	24
Figure 3.6	Schéma de la discrétisation de différentes méthodes de différences finies	26
Figure 3.7	Propriétés thermiques apprises par le MCE	30
Figure 4.1	Température au centre des trois couches principales d'un ruban uniforme	33
Figure 4.2	Schéma 3-D d'un ruban HTS : apprentissage le long de $A - B$ au centre de la couche	34
Figure 4.3	Schéma de la stratégie suggérée afin de remplacer la MEF	35
Figure 4.4	Schéma d'une géométrie 2-D d'un ruban HTS uniforme	36
Figure 4.5	Propriétés thermiques et électriques utilisées pour le modèle électro- thermique	37
Figure 4.6	Exemple d'un profil d'inhomogénéités $\Phi(x)$ pour $n_D = 4$	40
Figure 4.7	Évolution de la température en fonction du temps pour deux régimes de stabilité	41
Figure 4.8	Rampes de courant utilisées	42
Figure 4.9	Schéma de la résolution séquentielle découplée du modèle de référence	45
Figure 4.10	Schéma d'une cellule RAPDi et son dépliage temporel	46
Figure 4.11	Schéma du modèle d'AP utilisé dans le modèle RAPDi	50
Figure 4.12	Schéma détaillé de la cellule récurrente RAPDi	51

Figure 4.13	Simulation d'un long ruban HTS à partir d'une concaténation spatiale	52
Figure 5.1	Modèle d'apprentissage profond avec attention (APA)	57
Figure 5.2	Courbes d'apprentissage et prédictions (rubans uniformes)	60
Figure 5.3	Température longitudinale de la couche HTS d'un ruban uniforme pour un régime dominé par les pertes Joule	62
Figure 5.4	Température longitudinale de la couche HTS pour un régime dominé par le refroidissement	64
Figure 5.5	Erreur euclidienne moyenne en température	65
Figure 5.6	Densité de courant au centre de la couche HTS d'un ruban uniforme .	67
Figure 5.7	Potentiel électrique au centre de la couche HTS d'un ruban uniforme	68
Figure 5.8	Solution électrothermique d'un ruban uniforme de 1.4 mètre	70
Figure 6.1	Schéma 3-D d'un ruban CFD : les lignes $A - A'$ et $B - B'$	73
Figure 6.2	Vue de dessus de la température d'un ruban CFD lors d'un quench .	74
Figure 6.3	Température au centre de la couche HTS d'un ruban CFD 3-D	75
Figure 6.4	Longueurs de transfert d'un ruban CFD 3-D	77
Figure 6.5	Nuage de point : Correspondance $R_{inter,J}$ versus $I_{app}(t)$	80
Figure 6.6	Température longitudinale de la couche HTS d'un ruban CFD	89
Figure 6.7	Densité de courant au centre de la couche HTS	90
Figure 6.8	Potentiel électrique au centre de la couche HTS d'un ruban CFD . .	92
Figure C.1	Température longitudinale de la couche HTS d'un ruban CFD	108
Figure C.2	Densité de courant au centre de la couche HTS	109
Figure C.3	Potentiel électrique au centre de la couche HTS d'un ruban CFD . .	110

LISTE DES SIGLES ET ABRÉVIATIONS

Abréviations :

1-D	Une dimension
2-D	Deux dimensions
2G HTS CC	Second generation high temperature superconductive coated conductor - Supraconducteur recouvert de seconde génération à haute température critique
3-D	Trois dimensions
AC	Courant alternatif
ADAM	Méthode de l'estimation adaptative du moment
AP	Apprentissage profond
APA	Apprentissage profond avec attention
CFD	Current flow diverter - Divertisseur du flux de courant
CTL	Current transfer length - longueur de transfert du courant
CNN	Convolutional neural network - Réseau de neurones convolutif
DeepONet	Opérateur neuronal profond - Deep Operator Network
EDP	Équation aux dérivées partielles
EMQ	Énergie minimale de quench
GPU	Graphics Processing Unit - processeur graphique
GRU	Gated recurrent unit - unité récurrente à portes fermées
HTS	High temperature superconductor - Supraconducteur à haute température critique
LSTM	Long short-term memory - mémoire à court terme de longue durée
LTS	Low temperature superconductor - Supraconducteur à basse température critique
MCE	Modèle des coefficients effectifs
MEF	Méthode des éléments finis
MDF	Méthode des différences finies
NZPV	Normal zone propagation velocity - Vitesse de propagation de la zone normale
PINN	Réseau de neurones informés par la physique - Physics informed neural network
RAPDi	Récurrent d'apprentissage profond guidé par la physique basée sur les différences finies

REBCO	Rare-Earth Barrium Copper Oxide
RNN	Recurrent neural networks - Réseau de neurones récurrent
Typographie :	
$\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots$	Les variables en gras désignent un vecteur
$\ \cdot\ _2$	Norme euclidienne
$\cdot_\theta$	Quantité paramétrisée par les paramètres θ
u_i^n	Dénote la variable $u(x_i, t_n)$ où les indices i et n réfèrent à la discrétisation du maillage et du pas de temps respectivement
$\mathbf{T}_j^n$	Vecteur formé de la variable T_i^n et ses j plus proche voisin selon le maillage
$\langle e_X(t) \rangle$	Erreur euclidienne moyenne de la variable $X(t)$

Notations :

α_{NR}	Facteur d'amortissement dans l'algorithme de Newton-Raphson (s.d.)
Δx	Taille de la discrétisation du maillage (m)
Δt	Pas de temps (s)
ΔV_{rm}	Erreur relative maximale dans l'espace (s.d.)
κ	Conductivité thermique ($\text{W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$)
λ	Longueur de transfert du courant (m)
λ_R	Hyperparamètre de régularisation (s.d.)
Ω	Domaine (s.d.)
$\partial\Omega$	Frontière du domaine (s.d.)
$\Phi(x)$	Distribution spatiale des inhomogénéités (s.d.)
ρ	Résistivité électrique ($\Omega \cdot \text{m}$)
ρ_m	Densité massique ($\text{kg} \cdot \text{m}^{-3}$)
σ	Conductivité électrique ($\text{S} \cdot \text{m}^{-1}$)
σ_i	Largeur d'une inhomogénéité i (m)
θ	L'ensemble des paramètres d'un modèle d'AP (s.d.)
$a_\theta, b_\theta, c_\theta, d_\theta, e_\theta, f_\theta$	Préfacteurs du modèle RAPDi (s.d.)
$\mathbf{b}, \mathbf{c}$	Vecteur de biais – paramètres d'AP (s.d.)
d_k	Constante de normalisation pour le mécanisme d'attention (s.d.)
g	Modèle quelconque basé sur des lois physiques connues (s.d.)
h_i	Épaisseur de la couche i (m)
$\tilde{h}$	Coefficient de transfert thermique ($\text{W} \cdot \text{K}^{-1} \cdot \text{m}^{-2}$)
$\mathbf{h}_t$	Vecteur de l'état caché à l'instant t d'un RNN (s.d.)
l_s	Longueur d'une séquence (s.d.)

l_T	Longueur de diffusion thermique (m)
ℓ	Métrie d'erreur
n_D	Nombre d'inhomogénéités (s.d.)
n_0	Paramètre d'ajustement d'un matériau supraconducteur (s.d.)
$\mathbf{n}$	Vecteur unitaire normal (s.d.)
r	Résidu
$\mathbf{r}$	Vecteur des résidus (s.d.)
t	Temps (s)
t_f	Temps final (s)
v_{nz}	Vitesse de propagation de la zone normale ($\text{m}\cdot\text{s}^{-1}$)
w	Largeur du ruban (m)
w_R	Largeur de la résistance d'interface R_{inter} (m)
w_{CFD}	Largeur de la résistance CFD R_{CFD} (m)
x	Position longitudinale (m)
$x_{c,i}$	Position longitudinale d'une inhomogénéité (m)
$\mathbf{x}$	Vecteur d'entrée d'un réseau de neurones
$\mathbf{y}$	Vecteur de sortie d'un réseau de neurones
A_a	Section transversale de la couche stabilisatrice (m^2)
A_s	Section transversale de la couche supraconductrice (m^2)
$\bar{A}_i$	Amplitude d'une inhomogénéité (s.d.)
B	Champ magnétique (T)
B_c	Champ magnétique critique (T)
B_{c1}	Champ magnétique critique 1 pour un supraconducteur de type II (T)
B_{c2}	Champ magnétique critique 2 pour un supraconducteur de type II (T)
$\mathcal{B}$	Taille de lot (s.d.)
C	Capacité thermique combinée ($\text{J}\cdot\text{m}^{-2}\cdot\text{K}^{-1}$)
$\mathcal{C}$	Capacité thermique ($\text{J}\cdot\text{m}^{-3}\cdot\text{K}^{-1}$)
$\mathcal{C}_m$	Capacité thermique massique ($\text{J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$)
E	Champ électrique ($\text{V}\cdot\text{m}^{-1}$)
E_c	Champ électrique critique ($\text{V}\cdot\text{m}^{-1}$)
$\mathcal{F}$	Dynamique réelle d'un système étudié
$\mathcal{G}$	Opérateur solution
I	Courant (A)
I_{app}	Rampe de courant appliquée (A)
I_c	Courant critique (A)
I_{tr}	Thermal runaway current – courant d'emballement thermique (A)

J	Densité de courant ($\text{A}\cdot\text{m}^{-2}$)
$\bar{J}$	Densité de courant moyennée sur la section transverse du ruban ($\text{A}\cdot\text{m}^{-2}$)
$\mathbf{J}$	Densité de courant discrétisée ($\text{A}\cdot\text{m}^{-2}$)
J_c	Densité de courant critique ($\text{A}\cdot\text{m}^{-2}$)
J_{c0}	Densité de courant critique à T_0 ($\text{A}\cdot\text{m}^{-2}$)
$\mathcal{J}$	Matrice Jacobienne (s.d.)
K	Conductivité thermique effective ($\text{W}\cdot\text{K}^{-1}$)
L	Longueur du ruban (m)
$\mathcal{L}$	Fonction de perte
N	Nombre de données (s.d.)
N_x	Nombre de points selon le maillage (s.d.)
N_t	Nombre de points temporels (s.d.)
N_s	Nombre de simulations (s.d.)
Q	Chaleur dissipée ($\text{W}\cdot\text{m}^{-2}$)
$\tilde{Q}$	Pertes par effet Joule ($\text{W}\cdot\text{m}^{-3}$)
$R(\theta)$	Fonction de régularisation
R_{inter}	Résistance d'interface ($\Omega\cdot\text{m}^2$)
$R_{inter,T}$	Résistance d'interface effective du problème thermique ($\Omega\cdot\text{m}^2$)
$R_{inter,J}$	Résistance d'interface effective du problème électrique ($\Omega\cdot\text{m}^2$)
T	Température (K)
$\bar{T}$	Température moyennée sur la section transverse du ruban (K)
$\mathbf{T}$	Température discrétisée (K)
T_c	Température critique (K)
T_0	Température ambiante de l'environnement (K)
$U, \mathcal{V}, \mathcal{Q}, \mathcal{K}$	Matrices – paramètres d'AP (s.d.)
V	Potentiel électrique – Voltage (V)

LISTE DES ANNEXES

Annexe A	Discrétisation de l'équation de la chaleur par la méthode de Du Fort-Frankel	103
Annexe B	Algorithme de la méthode de Newton–Raphson pour l'équation du partage de courant	105
Annexe C	Solution électrothermique de ruban CFD où le modèle RAPDi n'est pas exact	107

CHAPITRE 1 INTRODUCTION

1.1 Mise en contexte

La découverte de la supraconductivité il y a maintenant plus de 110 ans a ultimement mené à la conception des rubans supraconducteurs d'aujourd'hui. Ces rubans possèdent des propriétés électromagnétiques hors pair, définies par leur résistivité nulle et leur diamagnétisme parfait. Ils sont alors une composante électrique idéale, capable de transporter jusqu'à des centaines d'ampères tout en produisant d'intenses champs magnétiques de plusieurs teslas sans la moindre génération de chaleur. Par ailleurs, leurs faibles dimensions leur confèrent des densités d'énergie élevées, pouvant ainsi réduire considérablement le poids des équipements électriques. Les nombreuses années d'études sur la supraconductivité ont fini par porter fruit : les matériaux supraconducteurs sont aujourd'hui la composante fondatrice de nombreuses technologies dont l'imagerie par résonance magnétique, les trains à sustentation magnétique (certaines technologies de Maglev), les câbles de transport à haute densité d'énergie, les moteurs légers et à haut rendement, les limiteurs de courant de court-circuit ou même dernièrement, les premières itérations de réacteurs à fusion nucléaire [1, 2]. La fusion nucléaire est décrite comme l'une des prochaines technologies révolutionnaires du 21^e siècle, tout juste après celle d'hier : l'intelligence artificielle.

Toutefois, l'enjeu principal de l'utilisation de rubans supraconducteurs est la possibilité d'apparition de « points chauds » dus à la variation locale du courant critique I_c le long du ruban, inhérente aux procédés de fabrication [3]. Ces points chauds peuvent engendrer un *quench*, c'est-à-dire un phénomène où l'entière du ruban transite brusquement de l'état supraconducteur à l'état normal (état résistif). Les quench peuvent parfois être à l'origine de dommages irréversibles à l'échelle locale, compromettant ainsi la fiabilité et la robustesse des technologies supraconductrices.

1.2 Problématique

La conception de rubans supraconducteurs nécessite alors des outils de modélisation fiables et précis. Ces outils permettent de mieux comprendre les phénomènes régissant la dynamique de quench, en plus de permettre de concevoir différentes architectures de rubans robustes face aux perturbations thermiques causées par les points chauds.

Toutefois, la modélisation des rubans supraconducteurs est une tâche ardue du point de vue numérique. Modéliser ces rubans implique nécessairement de résoudre des équations aux

dérivées partielles sur des domaines à deux ou trois dimensions, des maillages irréguliers ainsi que des problèmes à multi-échelles et multi-matériaux, dont plusieurs propriétés sont non-linéaires et varient de plusieurs ordres de grandeur. C’est notamment le cas de la résistivité électrique qui, à faible courant, peut mener à un comportement singulier lors de l’utilisation de méthodes numériques classiques [4]. Pour ces raisons, modéliser la dynamique de quench est intimement liée à des temps de simulation élevés, de nombreuses instabilités numériques ainsi qu’à des problèmes de convergence. La longueur des rubans, qui est en réalité de plusieurs mètres, est alors modélisée seulement sur quelques centimètres par souci de faisabilité.

Cependant, l’essor de l’intelligence artificielle durant les dernières années a permis le développement d’un tout nouvel outil de modélisation : l’apprentissage profond. Cet outil de modélisation a démontré, à plus d’une occasion, sa capacité à reproduire le comportement de phénomènes non linéaires pour une fraction du coût de calcul des méthodes numériques traditionnelles. Toutefois, malgré les avantages apparents, aucun modèle d’apprentissage profond n’a à ce jour été développé pour modéliser la dynamique de quench dans les rubans supraconducteurs. Il n’existe que des modèles quantifiant la probabilité d’occurrence d’un quench selon la géométrie du ruban.

1.3 Objectifs de recherche

L’objectif principal de ce mémoire est le développement d’un modèle numérique basé sur l’apprentissage profond afin de prédire de façon précise et peu coûteuse en temps de calcul la dynamique des points chauds dans les rubans supraconducteurs de grande longueur. Cet objectif se subdivise en 3 sous-objectifs :

1. *Concevoir un modèle mathématique à base d’apprentissage profond pour la dynamique du quench*

Ce mémoire est la continuité des travaux d’un post-doctorant de Polytechnique Montréal, dans lesquels un tout premier modèle a été développé pour la résolution de problèmes thermiques. L’objectif est alors de concevoir un modèle qui s’applique à la dynamique de quench à partir du modèle antérieur. Le modèle numérique doit être général pour une grande famille de conditions initiales, d’inhomogénéités du courant critique, ainsi que divers paramètres géométriques.

2. *Évaluer la précision du modèle en comparant ses résultats à ceux de modèles numériques validés expérimentalement*

Le modèle basé sur l’apprentissage profond doit être comparé à tous les stades du développement aux modèles déjà validés. Cette vérification permet d’identifier la qualité

de la modélisation tout en permettant d’identifier les faiblesses de l’approche. Des stratégies peuvent ainsi être formulées afin d’adresser les inconvénients du modèle proposé.

3. *Valider les prédictions du modèle sur des rubans de grande longueur*

Lorsque le modèle d’apprentissage profond est vérifié, des prédictions sur de longs rubans peuvent être effectuées afin de caractériser les performances de celui-ci. Il est toutefois essentiel d’établir une stratégie afin d’obtenir de telles simulations puisque les méthodes classiques de simulations par éléments finis sont impraticables.

1.4 Plan du mémoire

Ce mémoire est constitué de 7 chapitres. Le premier chapitre se veut une introduction au domaine et une brève présentation de la recherche. Les deuxième et troisième chapitres exposent les notions de bases et les fondements des rubans supraconducteurs ainsi que des méthodes d’apprentissage profond appliquées à la physique respectivement. Le quatrième chapitre présente la stratégie pour accélérer les simulations ainsi que la formulation du modèle d’apprentissage profond proposé. Le cinquième chapitre démontre la validité de l’approche pour des simulations 2-D et l’atteinte des objectifs de recherche. Le sixième chapitre va plus encore loin en démontrant la généralisation de l’approche pour des simulations 3-D. Finalement, le dernier chapitre est une conclusion et un sommaire du mémoire.

1.5 Contributions de ce mémoire

Ce mémoire introduit le tout premier modèle 1-D à base d’apprentissage profond capable de reproduire la dynamique complète du quench dans les rubans supraconducteurs pour des géométries 1-D, 2-D et 3-D. Les premiers résultats obtenus ont mené à une présentation invitée en juin 2024 dans le cadre du *9th International Workshop on Numerical Modelling of High Temperature Superconductors*. Un article de revue a également été rédigé à l’aide des résultats issus de ce mémoire. La publication est prévue après le dépôt du mémoire. Ces mêmes résultats ont aussi été présentés par affiche lors de la grande conférence du regroupement québécois des matériaux de pointes en mai 2025.

CHAPITRE 2 LES RUBANS SUPRACONDUCTEURS

2.1 La supraconductivité : définitions et propriétés

Cette sous-section a pour but de présenter les définitions formelles d'un supraconducteur, les interdépendances entre la température, le champ magnétique et la densité de courant, les divers types de supraconducteurs et leur comportement résistif.

2.1.1 Les paramètres critiques

La supraconductivité est souvent définie comme un état thermodynamique pour lequel la résistivité électrique est nulle ($\rho = 0$) à condition que la température T soit inférieure à la température critique T_c . Cette définition est exacte, mais toutefois incomplète puisqu'il est alors impossible de distinguer un supraconducteur d'un conducteur parfait.

En ce sens, une seconde propriété faisant souvent office de définition plus formelle est attribuée à l'état supraconducteur : l'effet Meissner-Ochsenfeld [5]. Découvert en 1933, cet effet consiste en l'exclusion de tout champ magnétique à l'intérieur du supraconducteur (état Meissner) par le biais de courants de surface. Il existe cependant une limite maximale de champ magnétique B qui peut être exclu, nommé le champ magnétique critique B_c . Au-delà de cette limite, le matériau n'est plus en mesure d'exclure le champ magnétique externe et perd alors l'état supraconducteur.

Par ailleurs, un troisième paramètre critique est attribué à l'état supraconducteur : la densité de courant critique J_c . Cette dernière est définie comme la valeur pour laquelle la densité de courant J fait disparaître l'état supraconducteur. Toutefois, une définition alternative est couramment utilisée en pratique et sera présentée dans la section 2.1.5.

2.1.2 Surface critique

La supraconductivité peut alors être vue comme un état de conduction parfaite, mais également de diamagnétisme parfait. Les paramètres critiques forment ainsi une frontière dans l'espace (T, B, J) délimitant l'état non dissipatif de l'état dissipatif. Cette frontière se nomme la surface critique et celle-ci est illustrée la figure 2.1. Par ailleurs, notons que comme les paramètres critiques dépendent les uns des autres, la surface critique peut s'avérer être de forme complexe.

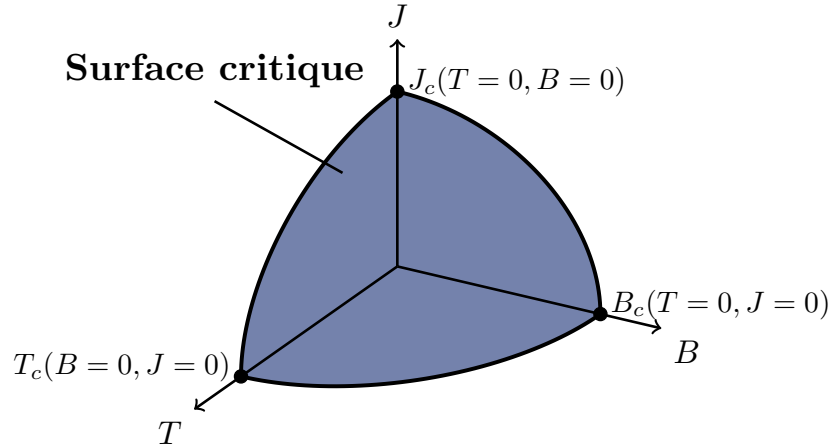


FIGURE 2.1 Surface critique d'un supraconducteur dans l'espace (T, B, J) .

2.1.3 Les LTS et HTS

On distingue deux catégories de supraconducteurs en fonction de leur température critique : les supraconducteurs à basse température critique (LTS) et les supraconducteurs à haute température critique (HTS). Découvert en 1911, les LTS sont caractérisées par une température critique inférieure à ~ 30 K, et requièrent un refroidissement cryogénique par liquéfaction de l'hélium liquide. De l'autre côté, les HTS ont été découverts en 1986 et regroupent les supraconducteurs possédant une température critique supérieure à 77 K. Plusieurs types de HTS sont refroidis à partir d'azote liquide, un refroidissement moins coûteux que le refroidissement par liquéfaction d'hélium [6, 7].

2.1.4 Les supraconducteurs de type I et II

Toutefois, classifier les supraconducteurs par leur température critique est inadéquat lorsqu'on étudie leurs propriétés électromagnétiques. On définit plutôt deux types de supraconducteurs : les type I et II.

Les supraconducteurs de type I sont caractérisés par une transition très abrupte, voire discontinue, entre l'état Meissner et l'état normal. Les supraconducteurs de type II sont quant à eux caractérisés par une transition raide, mais continue entre l'état Meissner et l'état normal, et possèdent deux champs critiques B_{c1} et B_{c2} au lieu d'un seul. Toutefois, la particularité principale des types II est la présence d'un état mixte lorsque le champ magnétique appliqué est compris entre B_{c1} et B_{c2} . Dans cet état, l'état Meissner coexiste avec des vortex laissant

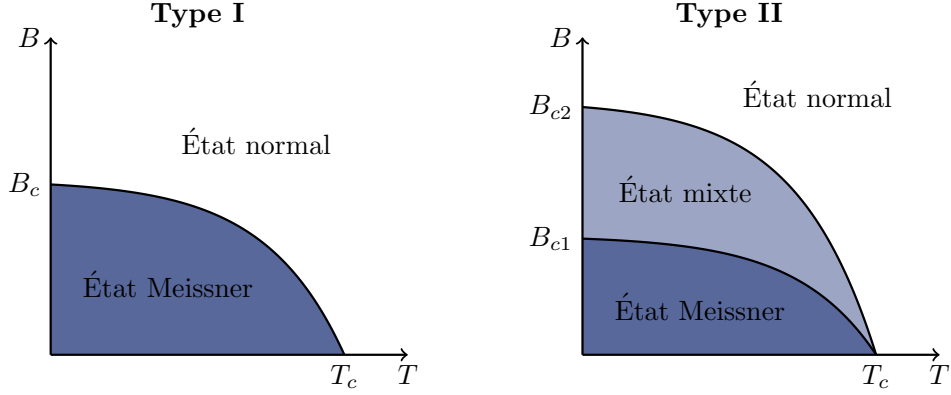


FIGURE 2.2 Comparaison des supraconducteurs de type I et type II dans l'espace (T, B) .

passer un quanta de flux magnétique en leur centre, ce qui résulte en un diamagnétisme imparfait. La figure 2.2 illustre alors la nuance entre les deux types de supraconducteurs.

Dans le cadre de ce mémoire, tous les supraconducteurs modélisés seront des REBCO (*Rare-Earth Barrium Copper Oxide*), un HTS de type II. Les REBCO sont couramment utilisés lors d'applications en raison de leur capacité à conserver un J_c élevé en présence de champs magnétiques intenses en plus de posséder un B_{c2} bien plus élevé que les B_c des type I.

2.1.5 La relation $E - J$ des HTS

La brusque variation de résistivité entre l'état supraconducteur et l'état normal induit une relation non linéaire entre le champ électrique E et la densité de courant J . En fonction de l'intensité de J , on peut définir trois régimes liés aux conditions microscopiques pour les HTS : le régime *flux pinning*, *flux creep* et *flux flow*. Ceci dit, il n'existe pas à ce jour de théories microscopiques explicatives du comportement exact de la résistivité dans ces trois régimes transitoires. Toutefois, il a été démontré que la relation $E - J$ peut être exprimée empiriquement par une loi de puissance [8] telle que

$$\mathbf{E} = E_c \left| \frac{\mathbf{J}}{J_c} \right|^n = \rho_s(\mathbf{J}) \mathbf{J}, \quad (2.1)$$

où $E_c = 1 \text{ } \mu\text{V/cm}$ est le champ électrique critique, n un paramètre empirique et ρ_s la résistivité du supraconducteur. En pratique, la littérature s'accorde à définir J_c telle que la densité de courant J qui produit le champ électrique E_c . Il s'agit d'une définition couramment utilisée dans le domaine expérimental et qui remplace souvent les définitions les plus formelles. Notons qu'il existe également d'autres lois empiriques afin de décrire la relation $E - J$, mais

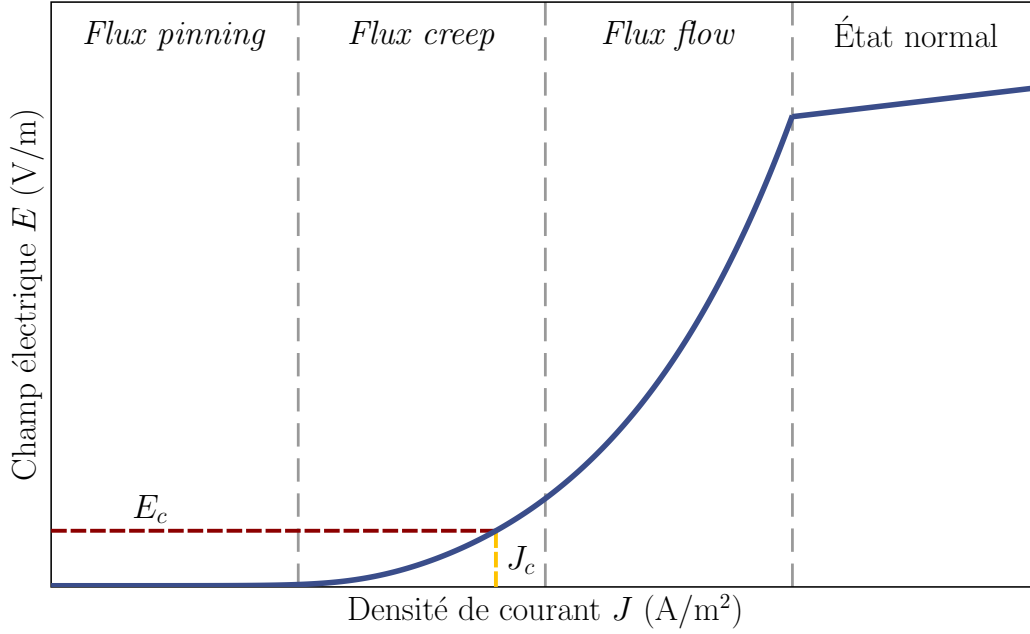


FIGURE 2.3 Relation non linéaire $E - J$ d'un supraconducteur de type II.

ces dernières ne seront pas abordées dans ce mémoire.

2.2 Les caractéristiques des rubans supraconducteurs

Les rubans supraconducteurs sont des HTS sous forme de fins rubans métalliques flexibles. Ce type de ruban est le fruit d'une longue recherche qui a mené aux *second generation high temperature superconductor coated conductor* ou 2G HTS CC. Les 2G HTS CC ont été conçus dans le but d'être robustes aux contraintes mécaniques en plus d'être résistants face aux instabilités thermiques. Ces propriétés sont obtenues par la déposition de diverses couches minces et forment alors la base fondatrice de plusieurs variétés d'architectures de rubans différentes. Afin de distinguer les principales architectures de 2G HTS CC, il est avant tout nécessaire d'aborder les caractéristiques des rubans supraconducteurs.

2.2.1 La vitesse de propagation de la zone normale (NZPV)

Le courant d'opération d'un ruban supraconducteur est près de son courant critique (I_c) dans des contextes d'applications électriques afin de maximiser la densité de courant utile sans détruire l'état supraconducteur. Cependant, un des grands problèmes des 2G HTS CC est la possibilité d'une variation locale du courant critique le long du ruban, inhérente au procédé de fabrication [3, 9].

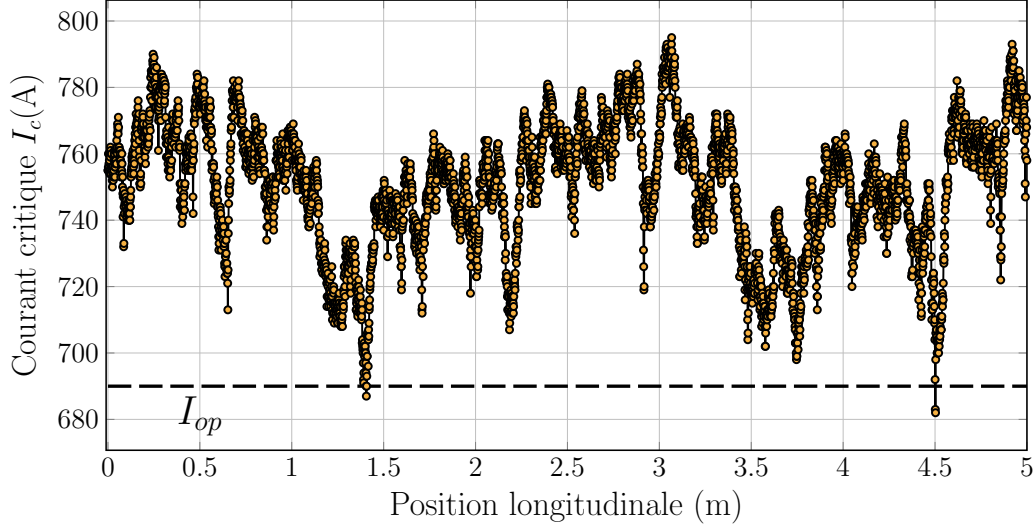


FIGURE 2.4 Distribution réelle de courant critique I_c à $T = 77$ K et $B = 0$ T le long d'un ruban supraconducteur et son courant d'opération I_{op} [10].

En ce sens, si le courant d'opération devient plus élevé que le courant critique localement, tel qu'illustré à la figure 2.4, l'état supraconducteur est perdu à l'échelle locale. Ces petites zones nouvellement résistives commencent alors à générer de la chaleur par effet Joule, engendrant des instabilités thermiques nommées points chauds. La température à l'emplacement des points chauds peut alors atteindre T_c , ce qui résulte en la création d'une zone à l'état normal appelée *zone normale*. Ces instabilités thermiques provoquent ainsi une élévation de la température environnante en raison de la dissipation de chaleur, et par conséquent, une croissance de la taille de la zone normale. En ce sens, le phénomène microscopique de point chaud peut provoquer un emballement thermique jusqu'à faire transiter brusquement l'intégralité du ruban supraconducteur à l'état normal. Cette transition est nommée *quench*. Ceci dit, les inhomogénéités du courant critique intrinsèques à la fabrication ne sont pas nécessairement l'unique facteur responsable de l'initialisation d'un quench. Des défauts ou faiblesses dans la distribution de I_c peuvent être engendrées par le biais d'un champ magnétique externe, d'une source de chaleur, de contraintes mécaniques, d'endommagements pendant la manipulation ou même par les pertes associées à un courant AC.

En conséquence, on définit pour les supraconducteurs la vitesse de propagation de la zone normale (NZPV). En l'occurrence d'un emballement thermique, la NZPV représente la vitesse longitudinale à laquelle la zone normale se propage et sera notée v_{nz} . La NZPV permet alors de déterminer la vitesse à laquelle un système peut détecter l'apparition d'un point chaud et il est donc avantageux de maximiser cette quantité. Ceci dit, la NZPV dépend toutefois d'une

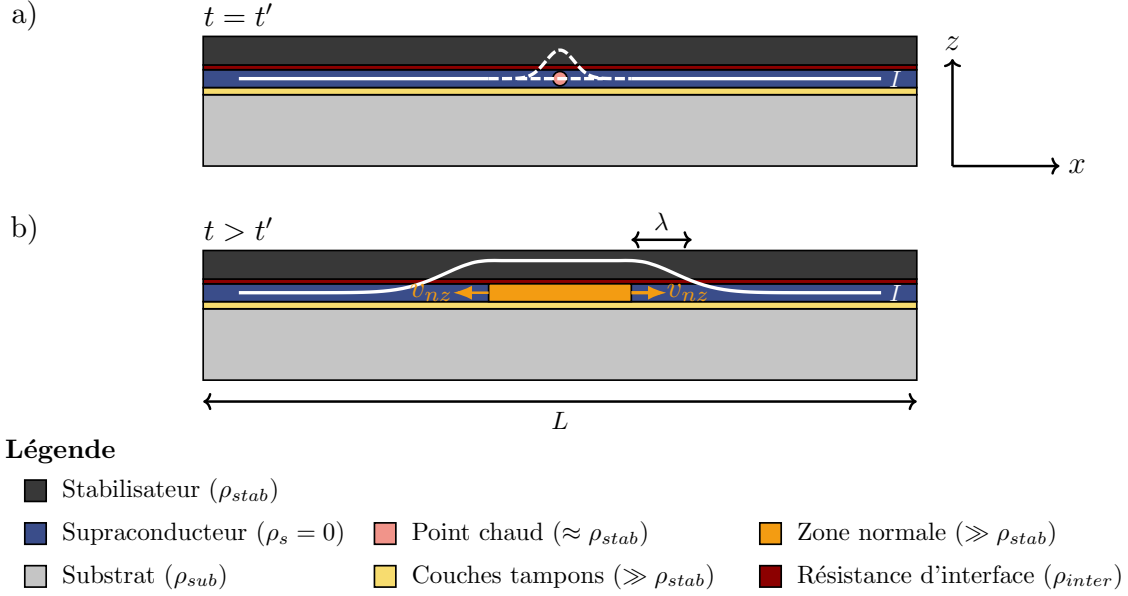


FIGURE 2.5 Vue longitudinale d'un partage de courant à l'intérieur d'un ruban HTS uniforme d'une longueur L durant (a) l'apparition d'un point chaud et (b) la propagation de la zone normale ultérieurement.

variété de facteurs dont la température, le courant d'opération, les propriétés électriques et thermiques des matériaux utilisés ainsi que l'architecture du ruban [11].

2.2.2 La longueur de transfert du courant (CTL)

Lors de la nucléation d'un quench le long d'un ruban HTS, un courant circulant dans la couche supraconductrice se voit contraint de dévier sa trajectoire vers la couche stabilisatrice selon une répartition dictée par les résistivités locales. Si les résistivités électriques des deux couches sont similaires, alors le courant se distribue partiellement entre celles-ci. Lors de la présence d'une zone normale, le courant circulera intégralement dans la couche stabilisatrice en raison de sa faible résistivité comparativement à celle des HTS de type céramique à l'état normal [12]. Ces deux scénarios de partage de courant sont illustrés à la figure 2.5. On définit ainsi la longueur de transfert λ ou la *current transfert length* (CTL) comme la distance nécessaire pour que le courant de transport transite intégralement de la couche supraconductrice vers la couche stabilisatrice. Mathématiquement, il est fréquent de définir λ comme la longueur pour laquelle l'amplitude du courant diminue d'un facteur e^{-1} .

2.2.3 La résistance d'interface

La longueur de transfert λ est proportionnelle à la résistance d'interface R_{inter} ($\Omega.m^2$) qui sépare les couches supraconductrices et stabilisatrices. La présence de cette interface résistive vient du contact électrique entre les deux couches qui, en pratique, peut varier énormément en fonction du procédé de fabrication du ruban et des matériaux utilisés [13]. Les travaux théoriques de Levin *et al.* ont démontré que la dépendance entre la NZPV et la longueur de transfert est quasi linéaire [14] et s'exprime telle que

$$v_{nz} \propto \max(\lambda, l_T), \quad (2.2)$$

où l_T est la longueur de diffusion thermique. Comme $\lambda \propto R_{inter}$, l'équation (2.2) indique alors qu'une importante résistance d'interface induit une NZPV élevée, facilitant ainsi la détection de quench avant qu'il n'y ait de dommages irréversibles. D'un autre côté, une faible résistance d'interface dissipe peu de chaleur, et de ce fait, minimise la probabilité qu'un quench survienne. La résistance d'interface doit donc être choisie de sorte à obtenir un équilibre entre les pertes ohmiques et la NZPV.

2.2.4 Stabilité thermique

Comme les rubans supraconducteurs sont continuellement exposés à un refroidissement cryogénique afin de conserver T_c , la présence de perturbations thermiques n'implique pas nécessairement l'apparition d'un emballement thermique. En fonction du bilan des flux de chaleur du système, on peut établir trois régimes.

Le régime de *cryostabilité* est lorsque la chaleur dissipée par effet Joule est inférieure à la chaleur évacuée par le refroidissement par convection de l'environnement cryogénique. Aucune zone normale ne peut alors exister sous de telles conditions. Le critère de Stekly est notamment un critère restrictif qui permet de vérifier ce bilan de chaleur [15].

Le régime *cold-end recovery* est lorsque la chaleur générée par effet Joule surpasse le refroidissement sans qu'un emballement thermique soit toutefois observé. Selon Maddok *et al.*, une zone normale peut se résorber à une certaine vitesse $-v_{nz}$ si les effets de conduction thermique sont dominants et que la taille de cette zone reste limitée [16]. On définit ainsi le courant d'emballement thermique I_{tr} comme le courant qui génère une zone normale plus grande que la taille limite. Dans le même ordre d'idées, l'énergie nécessaire pour créer cette zone de propagation minimale est nommée l'énergie minimale de quench (EMQ) [17, 18].

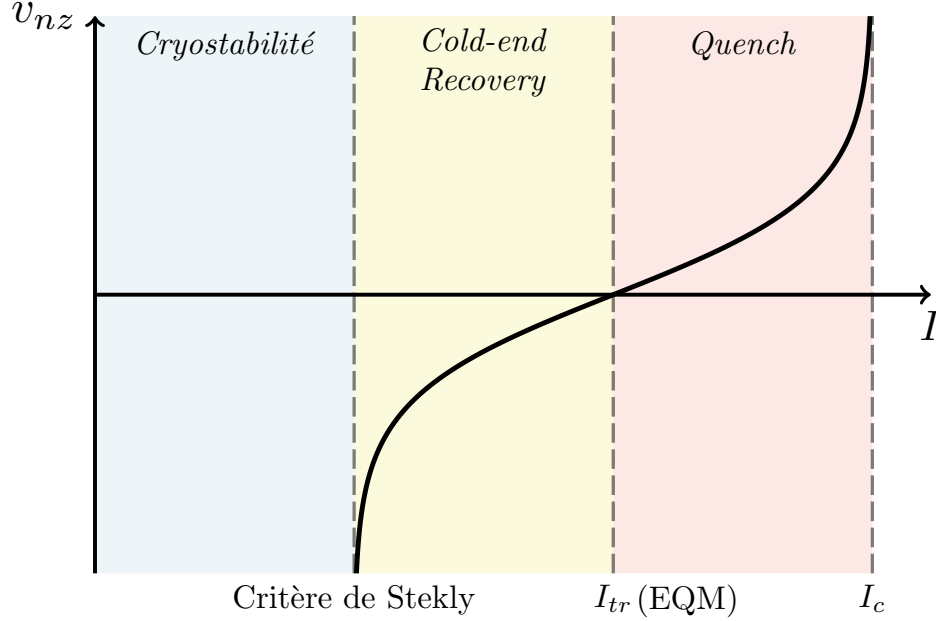


FIGURE 2.6 Représentation de la NZPV en fonction du courant I pour les 3 régimes de stabilité thermique. Adapté de [19].

2.3 Les architectures des rubans supraconducteurs

Le régime de *quench* survient alors lorsque la EMQ est dépassée. La figure 2.6 illustre les trois régimes en fonction de la NZPV et du courant I . Notons que la NZPV n'est pas définie pour $I > I_c$ puisque, dans une telle situation, l'intégralité du ruban transite simultanément à l'état normal.

Les rubans supraconducteurs sont conçus afin de préserver le comportement électromagnétique d'un HTS tout en assurant la stabilité thermique par le biais de couches de support et d'une NZPV élevée. En ce sens, les sous-sections suivantes présentent les détails de deux architectures de rubans HTS ainsi que leurs propriétés.

2.3.1 Architecture uniforme

L'architecture uniforme consiste en un empilement structuré d'un substrat, de couches tampons, d'une couche HTS, d'une interface résistive ainsi qu'une couche de stabilisation, le tout parfois enrobé d'une seconde couche stabilisatrice, tel que présenté à la figure 2.7a. Plus précisément, le substrat joue le rôle d'un support mécanique à la fois rigide et flexible en plus d'être un puits thermique. Les couches tampons servent quant à elles à protéger la couche HTS des contaminations potentielles du substrat tout en empêchant le courant d'être par-

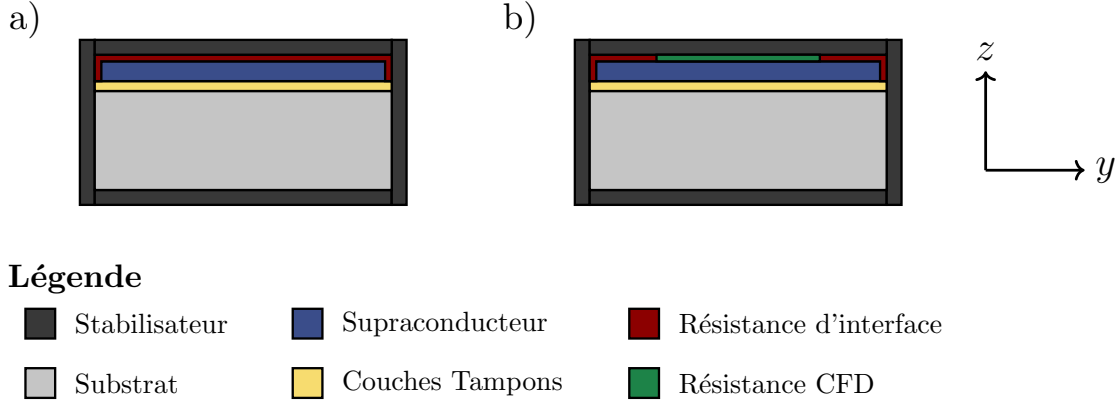


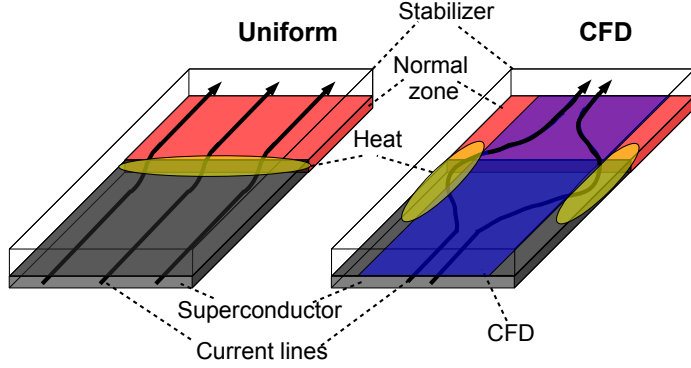
FIGURE 2.7 Vue transverse d'un ruban 2G HTS CC (a) uniforme enrobé et (b) CFD enrobé.

tager entre ces deux couches. La couche supraconductrice est responsable du transport du courant et la couche stabilisatrice agit à la fois comme masse thermique et comme chemin alternatif au courant en cas de quench. Ce chemin alternatif est plus adéquat à la propagation de chaleur en raison du milieu spécialement choisi pour sa résistivité électrique et sa conductivité thermique. Enfin, le ruban peut être enrobé à des fins de protection, mais également afin d'augmenter la NZPV. En contrepartie, enrober le ruban peut engendrer une déviation partielle du courant à l'intérieur du substrat [20]. Notons également que les dimensions d'un ruban HTS sont très particulières : l'épaisseur est de quelques dizaines de micromètres, la largeur de l'ordre des millimètres tandis que la longueur est de plusieurs mètres, voire kilomètres.

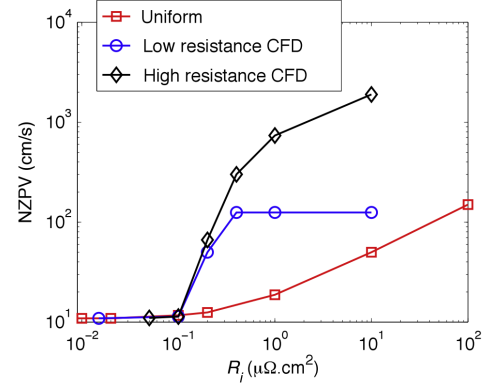
Ceci dit, les travaux expérimentaux de Lacroix *et al.* ont démontré que la NZPV peut être augmentée jusqu'à deux ordres de grandeur lorsque la résistance d'interface est augmentée de trois ordres de grandeurs, résultant en une NZPV de 0.05 à 10 cm/s [11, 21–23]. Cependant, l'inconvénient majeur de l'architecture uniforme reste l'importante chaleur générée par l'interface résistive qui devient rapidement un enjeu. Cette chaleur peut s'accumuler lors de l'injection du courant et aux extrémités de la zone normale.

2.3.2 Architecture CFD

En 2014, Lacroix et Sirois proposent une nouvelle architecture nommée le *Current Flow Diverter* (CFD) afin de pallier l'inconvénient majeur de l'architecture uniforme [24]. Pour ce faire, l'architecture CFD reprend l'intégralité de son prédécesseur tel qu'illustré à la figure 2.7b, à l'exception d'utiliser maintenant deux résistances d'interfaces : une résistance d'interface élevée (nommée résistance CFD) positionnée au centre du ruban, et une résistance



(a) Parcours du courant en 3-D. La zone bleue représente la résistance CFD.



(b) Relation entre la NZPV et la résistance d'interface.

FIGURE 2.8 Comparaison entre l'architecture uniforme et l'architecture CFD [24].

d'interface faible partout ailleurs.

La distribution du courant lors de l'occurrence d'un quench pour l'architecture CFD se distingue alors de celle de l'architecture uniforme telle qu'illustré à la figure 2.8a. Pour l'architecture CFD, le courant est dévié vers les extrémités du ruban lors de la présence d'une zone normale en raison de la résistance d'interface élevée au centre de celui-ci. Cette déviation du courant allonge alors la CTL, et par conséquent, induit une accélération de la NZPV d'après l'équation (2.2). Toutefois, la chaleur générée est relativement faible malgré l'accumulation de la densité de courant aux extrémités du ruban. L'architecture CFD propose ainsi la dynamique idéale : la CTL est majoritairement dictée par la résistance CFD tandis que les pertes par effet Joule sont déterminées principalement par la résistance d'interface faible. Ce phénomène de déviation du courant, d'où le nom CFD, permet donc résoudre l'inconvénient de l'architecture uniforme tout en augmentant la NZPV jusqu'à 50 m/s sous certaines conditions [20], soit plus d'un ordre de grandeur comparativement à l'architecture uniforme, tel que présenté à la figure 2.8b. Ceci dit, la NZPV reste proportionnelle à R_{inter} (résistance d'interface CFD), mais cette relation n'est plus quasi-linéaire, contrairement à l'architecture uniforme.

2.4 La modélisation des rubans 2G HTS CC

La modélisation numérique a démontré être un outil redoutable pour la conception d'architectures de ruban et l'analyse de leur comportement sous diverses excitations. Parmi les méthodes numériques les plus populaires, on retrouve la méthode des éléments finis (MEF),

des différences finies (MDF), des volumes finis ou même la méthode des moments [4]. Cependant, si l'on s'intéresse spécifiquement à la composante supraconductrice du système, il est commun de formuler le problème électromagnétique à l'aide de la MEF. Il est alors possible de résoudre le problème à partir des variables de champ (formulation $\mathbf{H}$ ou $\mathbf{E}$), ou en termes de potentiel (formulation $\mathbf{A} - V$ ou même $\mathbf{T} - \Omega$). Notons toutefois que certaines formulations s'expriment plus naturellement selon le problème qu'on cherche à résoudre.

Ceci dit, il est tout de même possible d'utiliser des formulations sans champ magnétique, à condition que les phénomènes thermiques soient dominants, comme c'est le cas lors de la propagation de la zone normale. Dans cette situation, il est raisonnable de conclure que la diffusion magnétique est bien plus rapide que la diffusion thermique, et par conséquent, que la diffusion thermique est le processus limitant [25]. Il est alors possible d'utiliser la formulation dite « électrothermique », qui résout conjointement les variables couplées de température T et de potentiel électrique V . Cette dernière formulation sera utilisée tout au long de ce mémoire.

2.4.1 Modèle 1-D pour la dynamique de quench

Le modèle électrothermique en géométrie 1-D proposé par Levin *et al.*, et validé expérimentalement, permet de modéliser la dynamique du profil de température ainsi que la distribution du courant entre la couche stabilisatrice et supraconductrice [14, 21, 26]. Comme le modèle de Levin *et al.* sera abondamment utilisé, celui-ci et ses limitations seront présentés dans les prochains paragraphes.

Considérons un ruban HTS uniforme, formé d'un empilement structuré d'un substrat (sub), d'une couche supraconductrice (s), d'une résistance d'interface R_{inter} et d'une couche stabilisatrice (a). Le modèle proposé par Levin *et al.* consiste à résoudre le système d'équations aux dérivées partielles

$$C(T) \frac{\partial T}{\partial t} - \frac{\partial}{\partial x} \left(K(T) \frac{\partial T}{\partial x} \right) = Q(x, T, J_s), \quad (2.3)$$

et

$$\frac{\partial}{\partial x} \left(h_a R_{inter} \frac{\partial J_a}{\partial x} \right) = \rho_a(T) J_a - \rho_s(x, T, J_s) J_s, \quad (2.4)$$

où

$$Q(x, T, J_s) = \underbrace{h_a \rho_a(T) J_a^2}_{\text{Chaleur } Q_a} + \underbrace{h_s \rho_s(x, T, J_s) J_s^2}_{\text{Chaleur } Q_s} + \underbrace{h_a^2 R_{inter} \left(\frac{\partial J_a}{\partial x} \right)^2}_{\text{Chaleur } Q_{inter}} + \underbrace{-2\tilde{h} \cdot (T - T_0)}_{\text{Chaleur } Q_{conv}}, \quad (2.5)$$

et où h_i est l'épaisseur de la couche i , $C(T) = \sum_i h_i C_i(T)$ est la capacité thermique combinée

et $\mathcal{C}_i(T)$ la capacité thermique, $K(T) = \sum_i h_i \kappa_i(T)$ est la conductivité thermique effective et $\kappa_i(T)$ la conductivité thermique, $\tilde{h}$ est le coefficient de transfert thermique, T_0 la température ambiante de l'environnement et i réfère aux indices du substrat, du supraconducteur et du stabilisateur respectivement. Les termes Q_a , Q_s et Q_{inter} représentent la chaleur surfacique générée par effet Joule au sein de la couche stabilisatrice, supraconductrice et de l'interface résistive respectivement, tandis que Q_{conv} représente la chaleur dissipée par le refroidissement cryogénique sur les deux plus grandes surfaces du ruban.

Ainsi, l'équation thermique (2.3) correspond à l'équation de la chaleur non linéaire décrivant l'évolution de la température selon des propriétés thermiques homogénéisées en W/m^2 , alors que l'équation électrique (2.4) décrit le partage de courant entre les couches supraconductrice et stabilisatrice en présence d'une résistance d'interface. L'équation thermique est donc couplée à l'équation électrique par le terme de chaleur (2.5). Le modèle de Levin *et al.* suppose également que le courant est uniformément réparti à l'intérieur de chaque couche. Le courant appliqué $I_{app}(t)$ s'exprime alors par la loi de conservation

$$I_{app}(t) = I_a(x) + I_s(x), \quad (2.6)$$

où $I_a(x)$ et $I_s(x)$ sont respectivement les courants de la couche stabilisatrice et supraconductrice. La résistivité du supraconducteur est quant à elle exprimée par une loi de puissance, i.e.

$$\rho_s(x, T, J_s) = \min \left(\frac{E_0}{J_c(x, T)} \left| \frac{J_s}{J_c(x, T)} \right|^{n-1}, \rho_{s,n} \right), \quad (2.7)$$

où $\rho_{s,n}$ est la résistivité du supraconducteur à l'état normal. Par ailleurs, on remarque que l'équation de partage du courant (2.4) devient une équation transcendante en raison de la résistivité non linéaire (2.7).

L'approche préconisée par les auteurs pour résoudre ce système consiste en une résolution séquentielle découplée à l'aide de la méthode des différences finies. En considérant la solution connue $T(x, t)$ pour un pas de temps Δt donné, l'équation transcendante (2.4) est résolue itérativement pour obtenir $J_s(x)$. La solution $J_a(x)$ est par la suite calculée à partir de la loi de conservation (2.6). À partir de ces solutions, il est alors possible d'obtenir les champs électriques $E_s(x)$ et $E_a(x)$, ainsi que les potentiels électriques $V_s(x)$ et $V_a(x)$. Les termes $J_a(x)$, $J_s(x)$ et $T(x, t)$ sont alors utilisés pour calculer la chaleur générée $Q(x, T, J_s)$. L'équation thermique peut alors être résolue afin d'obtenir la solution au prochain pas de temps, i.e. $T(x, t + \Delta t)$. Notons que le courant $I_{app}(t)$ et les propriétés thermiques sont considérées constantes pour un Δt donné. Il en va de même pour les champs électriques et les profils de tension, qui sont des variables dérivées de $I_{app}(t)$.

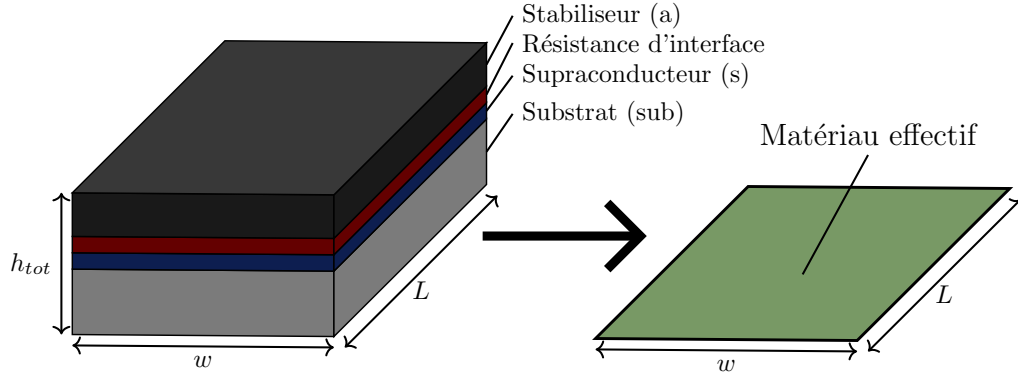


FIGURE 2.9 Gauche : Exemple d'un ruban supraconducteur uniforme 3-D en W/m^3 (pas à l'échelle). Droite : Ruban équivalent selon le modèle de Levin *et al.* avec des propriétés thermiques effectives en W/m^2 . h_{tot} dénote l'épaisseur totale du ruban.

2.4.2 Limitations du modèle analytique 1-D

Le modèle 1-D développé par Levin *et al.* décrit bien la physique inhérente à la dynamique de quench. Cependant, il présente deux grandes limitations, soit

1. Il repose sur l'hypothèse d'homogénéisation.
2. Il est basé sur une architecture uniforme, donc limité à cette géométrie (2-D).

L'homogénéisation (ou moyennage) des propriétés thermiques a l'avantage de permettre de représenter un modèle 2-D ou 3-D par un modèle 1-D. Toutefois, les travaux de Levin *et al.* ont démontré qu'une telle approximation n'est valide que si la chaleur est uniformément distribuée selon l'épaisseur du ruban, ou plus précisément, si la chaleur surfacique n'excède pas largement $10^4 W/m^2$ [14]. Si ce critère est respecté, la variation de température selon l'épaisseur n'est alors qu'une fraction de degré, soit très inférieure à la variation de température le long du ruban. Ainsi, lorsqu'on considère des échelles de temps de quelques secondes, telles que dans les problèmes d'aimants HTS, l'hypothèse d'homogénéisation est inexacte, mais somme tout acceptable. Toutefois, pour des échelles de temps des microsecondes, l'hypothèse d'homogénéisation est inadéquate en raison de la lente diffusion thermique selon l'épaisseur. Comme le substrat joue le rôle d'un puits thermique, un gradient se forme, rendant la distribution de chaleur non homogène. Le modèle de Levin *et al.* devient alors de moins en moins précis.

La seconde limitation du modèle analytique 1-D est qu'il ne peut pas être utilisé pour représenter des architectures CFD, puisque la distribution du courant est alors non homogène le long de la largeur (voir figure 2.8a). Le modèle de Levin *et al.* est alors incapable de décrire la dynamique manquante à partir d'une seule dimension géométrique.

CHAPITRE 3 TECHNIQUES DE MODÉLISATION : DE L'APPRENTISSAGE PROFOND À LA PHYSIQUE

Ce chapitre présente une introduction aux composantes principales de l'apprentissage profond (AP). Une fois ces outils introduits, le concept d'apprentissage profond guidé par la physique est présenté. Enfin, les travaux antérieurs sur lesquels repose ce mémoire sont discutés, suivis d'une brève présentation des problèmes de supraconductivité abordés par l'AP.

3.1 Notions de base en apprentissage profond

3.1.1 Les réseaux de neurones

Les réseaux de neurones sont un outil mathématique inspiré à l'origine du fonctionnement des neurones biologiques. Ils sont formés de trois types de couches : la couche d'entrée, les couches cachées et la couche de sortie. Chacune de ces couches possède un nombre de neurones ou unités, et chaque neurone reçoit en entrée un scalaire. De plus, chaque neurone est relié aux couches voisines par le biais de connexions pondérées nommées poids. Un schéma d'un réseau de neurones est présenté à la figure 3.1.

Mathématiquement, les réseaux de neurones sont représentés par une application séquentielle de transformations linéaires et de fonctions non linéaires sur un vecteur d'entrée $\mathbf{x}$. Les poids sont modélisés par des transformations linéaires W , tandis que les « fonctions d'activation » telles que $\text{ReLU}(\mathbf{x})$, $\tanh(\mathbf{x})$ ou $\text{sigmoid}(\mathbf{x})$ sont responsables de la non-linéarité du système. En ce sens, on peut exprimer un réseau de neurones de N couches tel que

$$\begin{aligned} \mathbf{z}^{(1)} &= \text{act}(W^{(1)}\mathbf{x} + \mathbf{b}^{(1)}), \\ \mathbf{z}^{(2)} &= \text{act}(W^{(2)}\mathbf{z}^{(1)} + \mathbf{b}^{(2)}), \\ &\vdots \\ \mathbf{y} &= \text{act}(W^{(N)}\mathbf{z}^{(N-1)} + \mathbf{b}^{(N)}), \end{aligned} \tag{3.1}$$

où $W^{(i)}$ est la matrice de poids i , $\mathbf{b}^{(i)}$ est le vecteur de biais i , $\text{act}(\cdot)$ est la fonction d'activation et $\mathbf{y}$ est le vecteur de sortie. Également, les paramètres θ d'un réseau de neurones correspondent à l'ensemble des valeurs qui forment les matrices de poids et les vecteurs de biais. On désigne par « apprentissage » ou « entraînement » le processus d'optimisation de ces paramètres.

Les réseaux de neurones sont également des approximateurs universels, c'est-à-dire qu'ils sont

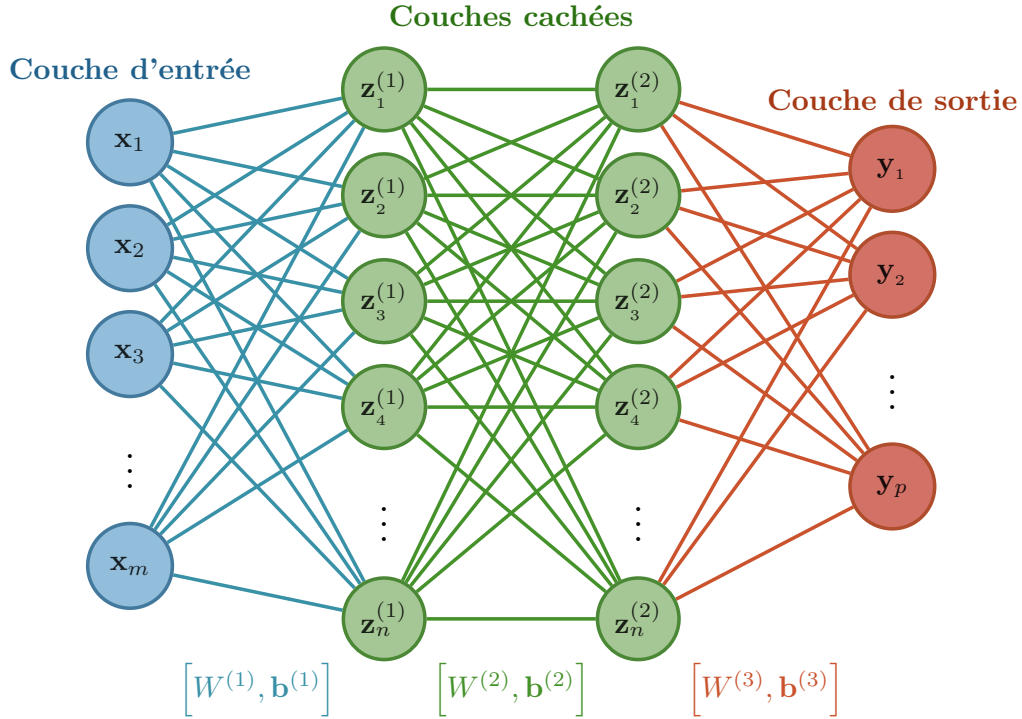


FIGURE 3.1 Schéma d'un réseau de neurones avec m neurones d'entrée, 2 couches cachées de n neurones et une couche de sortie de p neurones.

en mesure d'approximer n'importe quelle fonction continue à une précision arbitraire si le nombre de neurones est suffisant [27]. Toutefois, l'effervescence des réseaux de neurones est principalement expliquée par le développement des processeurs graphiques (GPU) au cours des dernières années. Les GPU permettent en effet de paralléliser les opérations matricielles comme celles des réseaux de neurones, permettant alors l'utilisation de très gros modèles d'AP en raison du temps d'entraînement et des coûts de calculs réduits par rapport à il y a une vingtaine d'années.

3.1.2 Les réseaux récurrents

Bien que les réseaux de neurones soient des approximateurs universels, ils ne sont pas toujours adaptés au traitement des données séquentielles telles que le langage naturel ou même les systèmes dynamiques. Pour ces problèmes, il est essentiel de tenir compte de l'état présent,

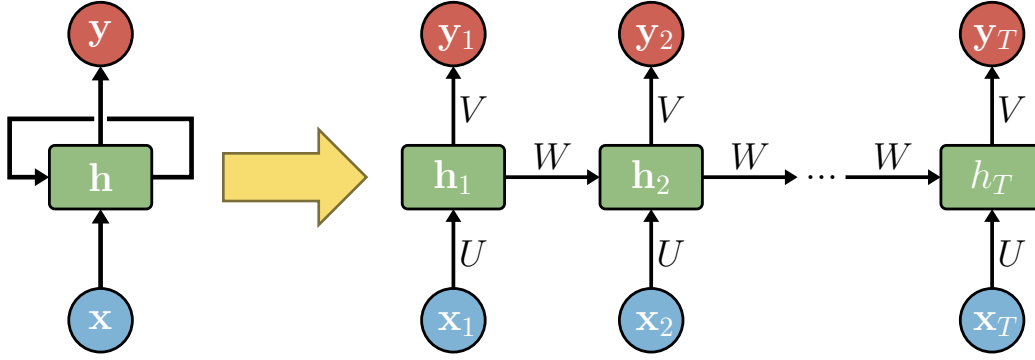


FIGURE 3.2 Schéma d'un réseau de neurones récurrent. Le schéma compact d'un RNN est illustré à gauche tandis que son dépliage temporel est illustré à droite.

mais également des états passés et leur dépendance temporelle respective. Afin de répondre à ce besoin, les réseaux de neurones récurrents (RNN) ont été introduits [28]. Les RNN peuvent être considérés comme des systèmes dynamiques à temps discret, spécialement adaptés pour de petites séries de temps. Pour une séquence d'entrée $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T$, les RNN calculent un état caché $\mathbf{h}_t$ à chaque instant t à partir du contexte présent $\mathbf{x}_t$ et passé $\mathbf{h}_{t-1}$. Cet état caché $\mathbf{h}_t$ peut être interprété comme un résumé condensé de l'état présent du système sous la forme d'un vecteur. La valeur de sortie $\mathbf{y}_t$ est quant à elle calculée à partir de $\mathbf{h}_t$. On modélise mathématiquement un RNN ainsi :

$$\begin{aligned} \mathbf{h}_t &= \tanh(U\mathbf{x}_t + W\mathbf{h}_{t-1} + \mathbf{b}), \\ \mathbf{y}_t &= \text{act}(\mathcal{V}\mathbf{h}_t + \mathbf{c}), \end{aligned} \tag{3.2}$$

où les paramètres à apprendre sont les matrices U, W et $\mathcal{V}$ et les vecteurs de biais $\mathbf{b}$ et $\mathbf{c}$. Le schéma d'une cellule RNN est présenté à la figure 3.2. Ceci dit, les RNN présentent des problèmes de gradient lors de l'apprentissage (*vanishing/exploding gradient*) si les séquences sont trop longues, empêchant alors de capturer les dépendances temporelles lointaines. Dans le but d'adresser ce problème, une variante des RNN nommée unité récurrente à portes fermées (GRU) introduit la notion de porte afin de contrôler le flux d'information [29]. Les *Long Short-Term Memory* (LSTM) ont également été proposés comme alternative plus performante aux GRU, en introduisant quant à eux des portes d'entrée, d'oubli et de sortie [28].

3.1.3 Le mécanisme d'attention

Les réseaux récurrents sont toutefois limités par leur traitement séquentiel des données, empêchant alors leur application aux ensembles de données massifs. Pour cette raison, des

chercheurs de Google ont introduit en 2017 le mécanisme d'attention [30]. Contrairement aux RNN, GRU ou LSTM, le mécanisme d'attention est conçu afin d'être parallélisable en plus de permettre de capturer des dépendances temporelles lointaines.

L'objectif du mécanisme d'attention est de pondérer à l'aide d'un score les différentes parties d'une séquence afin de déterminer sur quelles parties le modèle doit se concentrer. Pour ce faire, on définit le mécanisme d'attention tel que

$$\text{Attention}(\mathcal{Q}, \mathcal{K}, \mathcal{V}) = \underbrace{\text{softmax}\left(\frac{\mathcal{Q}\mathcal{K}^T}{\sqrt{d_k}}\right)}_{\text{Score}} \mathcal{V} \quad (3.3)$$

où $\mathcal{Q} = W_{\mathcal{Q}}\mathbf{x}$ est la matrice de requêtes, $\mathcal{K} = W_{\mathcal{K}}\mathbf{x}$ est la matrice de clés, $\mathcal{V} = W_{\mathcal{V}}\mathbf{x}$ est la matrice de valeurs, W_i est une transformation linéaire, $\mathbf{x}$ est le vecteur d'entrée et d_k est une constante de normalisation. Ce mécanisme parallélisable vient toutefois avec un coût : il ne tient pas compte de l'ordre naturel de la séquence. Afin de régler ce problème, un encodage positionnel est appliqué sur $\mathbf{x}$, permettant alors de représenter une position relative au sein d'une séquence. Notons qu'il est possible d'utiliser plusieurs mécanismes d'attention en parallèle afin d'augmenter la généralisation d'un modèle. C'est ce qu'on nomme le mécanisme d'attention multi-têtes (*Multi-Head Attention*). De plus, le mécanisme d'attention ne se limite pas qu'aux données séquentielles : il peut tout autant capturer des dépendances spatiales, comme c'est le cas en vision par ordinateur [31].

3.1.4 L'entraînement de modèles

L'entraînement d'un modèle d'apprentissage profond peut être formulé tel que l'on cherche l'ensemble de paramètres θ qui minimise la fonction de perte $\mathcal{L}$ (ou fonction objectif), qui quantifie l'erreur commise par le modèle pour une tâche donnée. Le problème d'optimisation s'exprime donc tel que

$$\min_{\theta} \mathcal{L}(\theta) \quad \text{avec} \quad \mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N \ell(y_i, u_{\theta}(\mathbf{x}_i)), \quad (3.4)$$

où l'on considère l'ensemble de données $\{\mathbf{x}_i, y_i\}_{i=1}^N$, où N est le nombre de données, ℓ est la métrique d'erreur et $u_{\theta}(\mathbf{x}_i)$ est la prédiction du réseau de neurones. Notons que l'indice θ signifie que la quantité considérée est paramétrisée par les paramètres θ . Ceci dit, comme le problème d'optimisation n'est pas convexe en raison des fonctions d'activation et de la multiplication successive des matrices de poids, l'équation (3.4) est résolue à partir d'algorithmes itératifs. Les algorithmes itératifs calculent la rétropropagation du gradient de $\mathcal{L}(\theta)$ afin de

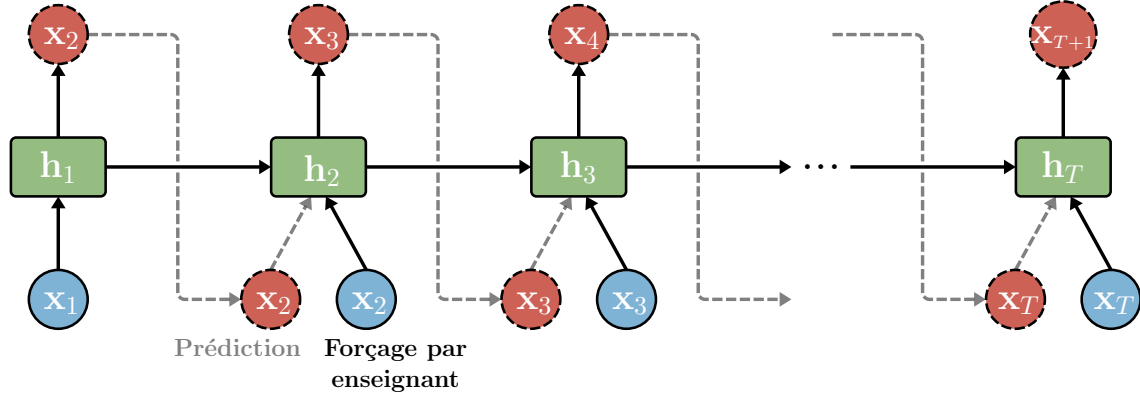


FIGURE 3.3 Schéma d'un RNN récurrent entraîné à l'aide du forçage par enseignant. Les lignes pointillées représentent la prédiction du RNN. Durant l'entraînement, on utilise les valeurs d'entrées en bleu.

déterminer la correction à apporter sur chacun des paramètres θ . Pour les méthodes d'ordre 1, la descente de gradient stochastique et la méthode de l'estimation adaptative du moment (ADAM) sont parmi les algorithmes les plus populaires de leur catégorie. Il existe également des méthodes d'ordre 2 telles que l'algorithme L-BFGS (quasi-Newton), qui consiste à approximer la matrice Hessienne. Toutefois, les algorithmes d'ordre 1 sont souvent privilégiés en raison de leur faible coût de calcul.

Notons également que la convergence du processus d'apprentissage peut être accélérée par l'ajout de normalisation par lot ou de normalisation par couche au sein de l'architecture des modèles d'AP [32,33]. De plus, les problèmes de disparition de gradient propre aux réseaux de neurones très profonds peuvent être résolus à l'aide de connexions résiduelles [34].

Par ailleurs, les modèles récurrents tels que certains types de RNN sont susceptibles d'être exigeants en calcul ou difficiles à faire converger lors de l'entraînement avec de longues séquences. Le forçage par enseignant est alors une technique d'apprentissage à faible coût pouvant remédier à ces problèmes. Le forçage par enseignant consiste à fournir la prédiction exacte à chaque instant t au lieu d'utiliser la prédiction du RNN lors de l'entraînement. Cette technique peut également servir à pré-entraîner les modèles récurrents. Toutefois, l'inconvénient majeur du forçage par enseignant est que les erreurs vues durant l'entraînement sont relativement faibles comparativement à celles lors de l'inférence. Ainsi, lorsque le modèle prédit une longue séquence par lui-même, les erreurs s'accumulent et les prédictions se détériorent alors au fil de la séquence.

3.1.5 Le surapprentissage

Un entraînement prolongé sur un ensemble fini de données peut mener à l'effet de surapprentissage. Ce phénomène consiste à apprendre si fidèlement l'ensemble d'entraînement qu'on en vient à apprendre les spécificités et le bruit numérique des données, au détriment de la distribution réelle de celles-ci. Le surapprentissage compromet donc la généralisation des prédictions d'un modèle.

Afin de prévenir ce phénomène, on monitore les performances d'un modèle sur un ensemble indépendant à celui de l'entraînement, appelé l'ensemble de validation. L'arrêt hâtif consiste alors à arrêter l'entraînement lorsqu'un déclin des performances est observé sur l'ensemble de validation, ce qui permet de conserver la généralisation du modèle. Par ailleurs, le surapprentissage peut se traduire par des valeurs de paramètres excessivement grandes. La régularisation est donc une seconde technique de prévention qui consiste à imposer une contrainte souple au problème d'optimisation telle que

$$\min_{\theta} \left(\mathcal{L}(\theta) + \lambda_R R(\theta) \right), \quad (3.5)$$

où λ_R est un hyperparamètre et $R(\theta)$ est une fonction de régularisation. La régularisation L_2 ($R(\theta) = \|\theta\|_2^2$) est notamment un exemple de fonction de régularisation qui décourage les paramètres de grandes valeurs. Également, les réseaux de neurones peuvent souffrir de co-adaptation, un phénomène d'interdépendance excessive entre neurones. La technique du *dropout* permet d'éviter cette problématique en désactivant aléatoirement des neurones durant l'entraînement afin que l'entièreté des paramètres du réseau soit utilisée [35].

Par ailleurs, un réseau de neurones tel qu'illustré à la figure 3.1 peut être qualifié de réseau *feed forward*. Lorsque l'architecture d'un modèle d'AP se complexifie, les réseaux de neurones deviennent alors des composantes du modèle. Dans un tel cas, il est fréquent de schématiser le modèle par un schéma « en blocs » comme celui de la figure 3.4.

3.2 L'apprentissage profond guidé par la physique

La physique qui régit un système est régulièrement modélisée par un ensemble d'équations aux dérivées partielles (EDP). De nombreuses méthodes numériques permettent de résoudre ces EDP en discrétisant l'espace, et selon leur formulation, celles-ci peuvent même assurer la conservation de quantités physiques. Toutefois, ces méthodes s'apprêtent moins bien lorsque les EDP qui régissent le système étudié sont inconnues, en plus d'être souvent coûteuse en calcul en raison de la fine discrétisation nécessaire.

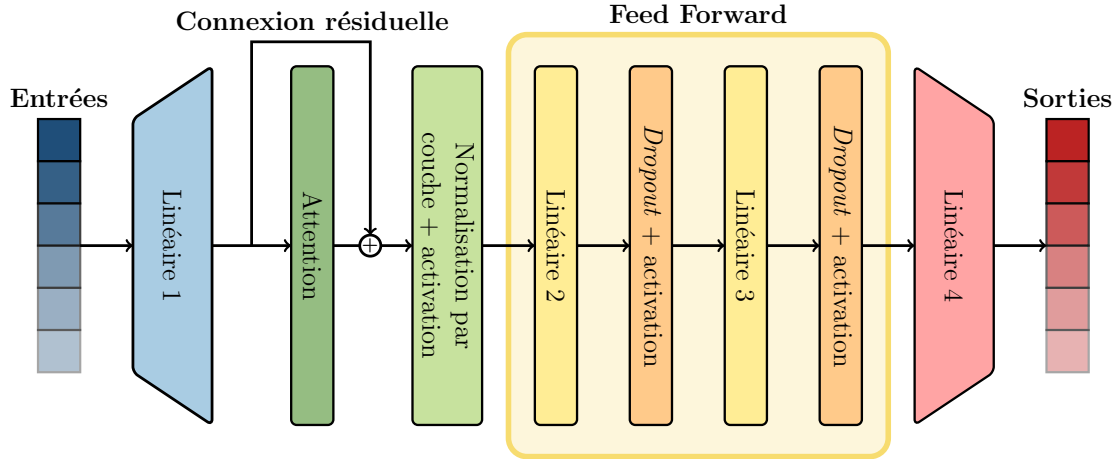


FIGURE 3.4 Schéma d'un modèle d'AP formé d'une connexion résiduelle, d'un mécanisme d'attention, d'une couche de normalisation, d'un réseau *feed forward* avec des couches de *dropout*. Les transformations linéaires sont notées « Linéaire i ». Les entrées et sorties sont des vecteurs.

Les méthodes d'AP de style boîte noire offrent alors une solution à ces problèmes en modélisant le système étudié par un réseau de neurones. Le modèle de boîte noire trouve une correspondance directe entre les entrées et sorties du système, permettant alors d'accélérer les calculs en évitant les intégrations numériques des méthodes traditionnelles. Comme ce type de modèle est purement guidé par les données, celui-ci peut représenter une dynamique inconnue avec un minimum d'hypothèses. Cependant, la modélisation par boîte noire s'est avérée être à maintes reprises ininterprétable, imprécise et incohérente en raison des lois de conservation violées [36].

Ceci dit, la modélisation par boîte noire a ouvert la porte aux modèles combinant la physique aux données, un domaine qu'on appelle aujourd'hui l'apprentissage profond guidé par la physique. Les modèles d'AP guidé par la physique introduisent les lois physiques par le biais de contraintes au sein de leur architecture. Plusieurs de ces modèles sont présentés dans les sous-sections suivantes. La figure 3.5 illustre une vue d'ensemble de ces modèles.

3.2.1 Les modèles hybrides

Les modèles hybrides cherchent à approximer la dynamique réelle $\mathcal{F}$ d'un système à partir de lois physiques connues [37]. Pour ce faire, on décompose la dynamique du système de telle sorte que $\mathcal{F} = g + r$, où g est un modèle basé sur les lois physiques connues et r un résidu décrivant la physique manquante. En ce sens, la majorité des lois physiques sont incorporées

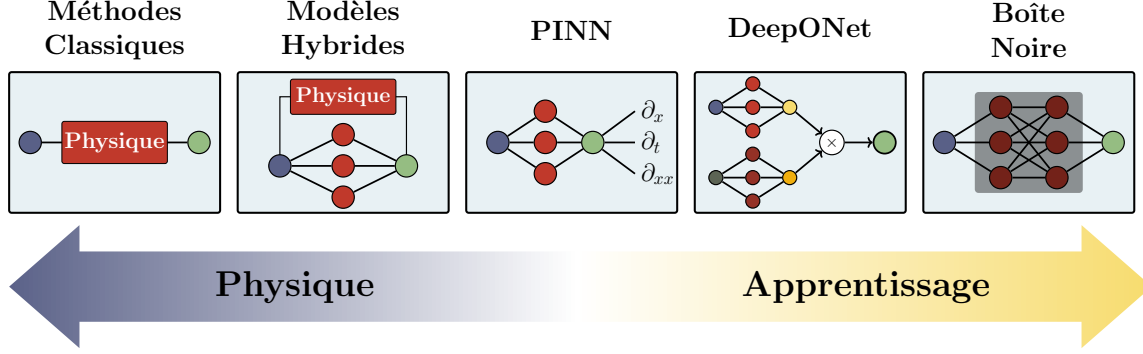


FIGURE 3.5 Schéma de différents modèles d'apprentissage profond guidé par la physique. Les modèles à gauche ont une dépendance plus importante à la physique, tandis que plus on se déplace à droite, plus les modèles dépendent de l'apprentissage. Figure inspirée de [37].

via g , un modèle résolu à partir de méthodes traditionnelles. Le résidu r est quant à lui approximé par un modèle d'AP. La fonction de perte s'écrit alors comme

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N \ell(r_i, u_\theta(\mathbf{x}_i)), \quad (3.6)$$

où $r_i = y_i - g(\mathbf{x}_i)$ est le résidu associé aux données $\mathbf{x}_i$ et $g(\mathbf{x}_i)$ est la prédiction du modèle classique. Ceci dit, l'hypothèse que $\mathcal{F}$ s'écrit sous forme additive implique que les coûts de calcul de $\mathcal{F}$ sont intrinsèquement reliés à ceux de g . D'ailleurs, selon cette formulation, la dynamique étudiée nécessite parfois d'être stable afin que le problème converge. En présence d'instabilités ou si $u_\theta(\mathbf{x}_i)$ semble trop flexible, des contraintes souples peuvent être ajoutées.

3.2.2 Les réseaux de neurones informés par la physique (PINN)

Les réseaux de neurones informés par la physique (PINN) consistent à résoudre des EDP à partir d'un réseau de neurones où les conditions frontières, les conditions initiales et les EDP sont implémentées sous forme de régularisation [38]. Les dérivées partielles de la solution sont quant à elles calculées de façon exacte à l'aide d'une différenciation automatique [39]. En ce sens, en posant $\mathcal{R}(u) = 0$ comme étant le résidu d'une EDP avec $u(x, t)$ la solution exacte et $u_\theta(x, t)$ la solution du réseau de neurones, on exprime la fonction de perte telle que

$$\mathcal{L}(\theta) = \underbrace{\lambda_d \frac{1}{N_d} \sum_{i=1}^{N_d} \left| \mathcal{R} \left[u_\theta(x_i, t_i) \right] \right|^2}_{\text{Résidu moyen de l'EDP}} + \underbrace{\lambda_f \frac{1}{N_f} \sum_{j=1}^{N_f} \left| u(x_j, t_j) - u_\theta(x_j, t_j) \right|^2}_{\text{Résidu moyen des conditions frontières et initiales}}, \quad (3.7)$$

où (x_i, t_i) et (x_j, t_j) sont les N_d points à l'intérieur du domaine et les N_f points sur les conditions frontières et initiales respectivement, alors que λ_d et λ_f sont des hyperparamètres. Ceci dit, les contraintes souples n'assurent pas nécessairement la conservation des lois physiques, mais servent plutôt à « guider » la solution du réseau de neurones vers une solution physique. Notons par ailleurs que les PINN sont indépendants du maillage lors de l'inférence. Par contre, ceux-ci doivent être réentraînés à chaque fois qu'on souhaite changer le maillage de la solution, ou lorsqu'au moins l'un des coefficients de l'EDP est varié. Il a également été démontré que les PINN classiques sont imprécis lorsque les coefficients de l'EDP sont non triviaux ou de grande amplitude [40].

3.2.3 Les opérateurs neuronaux profonds (DeepONet)

Les opérateurs neuronaux profonds (DeepONet) sont des modèles d'AP qui consistent à approximer l'opérateur solution d'une EDP paramétrique étudiée [41]. Si on considère les paramètres $\mathbf{s}$ et la solution de l'EDP $u(x, t)$, un DeepONet cherche à trouver un opérateur $\mathcal{G}$ tel que $u(x, t) = \mathcal{G}\{\mathbf{s}\}$. Pour ce faire, l'architecture du modèle est formée de deux réseaux de neurones liés, soit le réseau *branch* et le réseau *trunk*. Si on possède les couples de données $\{\mathbf{s}^{(i)}, x_i, t_i\}_{i=1}^N$, le réseau *branch* reçoit les paramètres $\mathbf{s}^{(i)}$ en entrée et produit le vecteur $\mathbf{b} \in \mathbb{R}^p$ en sortie, où p est le nombre de neurones de sortie, tandis que le réseau *trunk* reçoit les points (x_i, t_i) et génère le vecteur $\mathbf{t} \in \mathbb{R}^p$ en sortie. La fonction de perte s'écrit alors

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N \left| u(x_i, t_i) - \underbrace{\sum_{k=1}^p b_k(\mathbf{s}^{(i)}) t_k(x_i, t_i)}_{=\mathcal{G}_\theta\{\mathbf{s}\}(x,t)} \right|^2, \quad (3.8)$$

où l'indice k désigne le $k^{\text{ième}}$ élément du vecteur et $\mathcal{G}_\theta\{\mathbf{s}\}(x, t)$ est l'approximation de l'opérateur $\mathcal{G}$ par le réseau de neurones. Notons que les DeepONet sont indépendants du maillage, et contrairement aux PINNs, ils n'ont pas besoin d'être réentraînés lorsque les paramètres sont variés. Ceux-ci peuvent également, sous certaines conditions, changer la résolution du maillage sans perte de généralité, et ils ne requièrent pas de connaître l'EDP étudiée. Toutefois, les DeepONet ont de la difficulté à résoudre des problèmes multi-échelle ou à maillage irrégulier [42, 43].

3.3 Modèle hybride appliqué aux problèmes thermiques

Comme ce mémoire est la continuité des travaux post-doctoraux d'un ancien étudiant, cette section a pour guise de présenter rapidement la méthode numérique au coeur du modèle

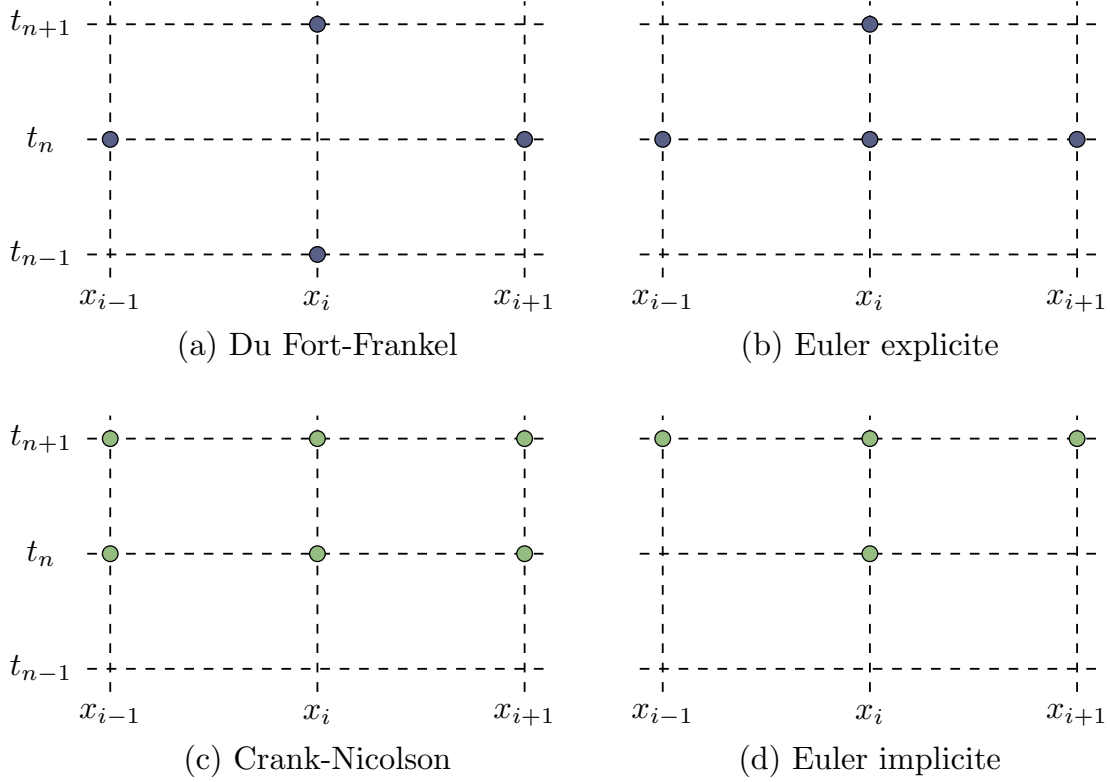


FIGURE 3.6 Schéma de la discrétisation de (a) Du Fort-Frankel, (b) Euler explicite, (c) Crank-Nicolson et (d) Euler implicite autour du point (x_i, t_n) par la méthode des différences finies. Les cercles représentent les degrés de liberté intervenant dans la discrétisation par différences finies. La couleur bleue indique les schémas explicites, tandis que la couleur verte correspond aux schémas implicites. Figure inspirée de [44].

hybride qui a été développé ainsi que le modèle en question.

3.3.1 La discrétisation de Du Fort-Frankel

Le schéma de Du Fort-Frankel, présenté à la figure 3.6a, est une discrétisation explicite et inconditionnellement stable servant à approximer des dérivées par la méthode des différences finies. Ce schéma est particulièrement adapté aux équations aux dérivées partielles paraboliques, dont l'équation de la chaleur. Concrètement, le schéma de Du Fort-Frankel consiste à approximer

$$\frac{\partial u_i^n}{\partial t} \approx \frac{u_i^{n+1} - u_i^{n-1}}{2\Delta t} \quad \text{et} \quad u_i^n \approx \frac{u_i^{n+1} + u_i^{n-1}}{2}, \quad (3.9)$$

avec $u_i^n = u(x_i, t_n)$ où i et n dénotent les indices de discrétisation selon le maillage et le pas de temps respectivement. La discrétisation de Du Fort-Frankel est une méthode explicite

d'ordre 2 en espace et d'ordre 1 en temps, pouvant être ainsi vue comme une amélioration de la méthode d'Euler explicite dans le même sens que la méthode de Crank-Nicolson est une amélioration de la méthode d'Euler implicite [44]. Notons toutefois que la discrétisation de Du Fort-Frankel implique connaître la solution aux temps n et $n - 1$.

3.3.2 Le modèle des coefficients effectifs (MCE)

Le modèle des coefficients effectifs (MCE) est un modèle hybride développé par Jean-Hughes Fournier-Lupien, ancien chercheur postdoctoral au laboratoire. Le modèle, non publié, est spécifiquement adapté aux problèmes thermiques à géométrie plane [44]. Ce modèle 1-D permet de résoudre l'EDP thermique de la forme

$$\mathcal{C}(T) \frac{\partial T}{\partial t} - \frac{\partial}{\partial x} \left(\kappa(T) \frac{\partial T}{\partial x} \right) = \tilde{Q}(x, t) - \frac{S}{V} \tilde{h}(T - T_0)(T - T_0), \quad (3.10)$$

où T est la température, $\tilde{Q}(x, t)$ est une source de chaleur de forme gaussienne (W/m^3), S et V sont l'aire exposée à l'environnement et le volume du matériau respectivement, et $\tilde{h}(T - T_0)$ est le coefficient de transfert convectif dépendant de la température. L'application du MCE consiste alors à discrétiser l'équation (3.10) à l'aide du schéma de Du Fort-Frankel et d'approximer les propriétés thermiques $(\mathcal{C}, \kappa, \tilde{h})$ par trois réseaux de neurones notés $(\mathcal{C}_\theta, \kappa_\theta, \tilde{h}_\theta)$ respectivement. En ce sens, on obtient l'équation discrétisée

$$T_{\theta,i}^{n+1} = \frac{1}{1 + A_\theta} \left[(1 - A_\theta) T_i^{n-1} + A_\theta (T_{i+1}^n + T_{i-1}^n) + B_\theta (T_{i+1}^n - T_{i-1}^n)^2 + \gamma_\theta \tilde{Q}_i^n - \tilde{h}_\theta (T_i^n - T_0)(T_i^n - T_0) \right], \quad (3.11)$$

où

$$A_\theta = \frac{2\kappa_\theta(T_i^n)\Delta t}{\mathcal{C}_\theta(T_i^n)\Delta x^2}, \quad B_\theta = \frac{\Delta t}{2\mathcal{C}_\theta(T_i^n)\Delta x^2} \frac{\partial \kappa_\theta(T_i^n)}{\partial T}, \quad \gamma_\theta = \frac{2\Delta t}{\mathcal{C}_\theta(T_i^n)}, \quad (3.12)$$

$T_{\theta,i}^{n+1} = T_\theta(x_i, t_n)$ est la prédiction du modèle et Δx désigne la taille de la discrétisation du maillage. Rappelons que l'indice θ est introduit afin de dénoter les dépendances de l'équation (3.11) et (3.12) avec les paramètres θ des réseaux de neurones.

Le MCE permet donc de modéliser des phénomènes bien plus complexes que les modèles seulement à base de réseaux de neurones puisque approximer les dérivées partielles de l'EDP par les différences finies simplifie grandement la tâche d'apprentissage. En ce sens, l'équation (3.11) permet d'incorporer une partie de la dynamique étudiée par le biais d'une contrainte physique dure, ce qui guide ainsi la solution vers une solution physique. Ceci dit, bien que

le MCE demande des coûts en calcul plus élevés qu’un simple réseau de neurones, la nature explicite de la discrétisation permet d’obtenir une solution à des coûts relativement faibles comparativement à des discrétisations implicites.

Notons également que le MCE se distingue de la définition d’un modèle hybride de la sous-section 3.2.1. Plus précisément, les modèles hybrides tels que présentés précédemment posent que $T(x_i, t_{n+1}) \approx g(x_i, t_n) + u_\theta(x_i, t_n)$, où le réseau de neurones $u_\theta(x_i, t_n)$ cherche à compenser l’erreur associée au modèle g . À titre d’exemple, le modèle g pourrait être un modèle de différences finies. Toutefois, le MCE n’utilise pas ses réseaux de neurones $(\mathcal{C}_\theta, \kappa_\theta, \tilde{h}_\theta)$ sous une forme additive, mais plutôt sous la forme $T(x_i, t_{n+1}) \approx g_\theta(x_i, t_n) = T_\theta(x_i, t_n)$ (équation (3.11)). Dans cette seconde version, les réseaux de neurones sont incorporés à l’équation discrétisée issue des différences finies (modèle g), paramétrant alors celle-ci.

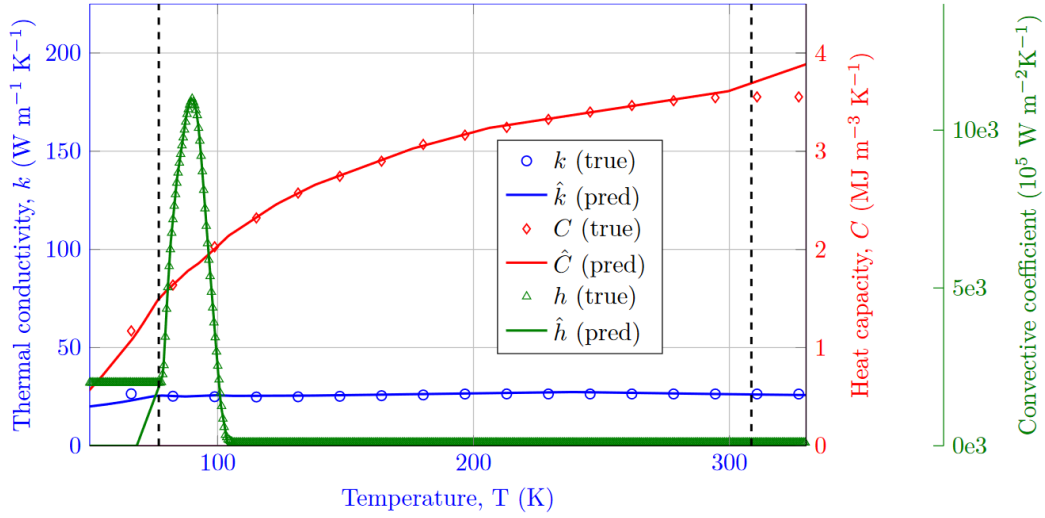
Les travaux de Fournier-Lupien *et al.* ont également démontré que le modèle des coefficients effectifs permet de résoudre l’équation thermique (3.10) pour une géométrie 1-D avec une précision supérieure à celle de la méthode des différences finies pour un maillage donné. Par ailleurs, non seulement le MCE est plus précis, mais une fois celui-ci entraîné, les réseaux de neurones qui le constituent $(\mathcal{C}_\theta(T), \kappa_\theta(T), \tilde{h}_\theta(T))$ convergent vers les coefficients thermiques réels de l’EDP (3.10) $(\mathcal{C}(T), \kappa(T), \tilde{h}(T))$, tels qu’illustrés à la figure 3.7a. Le modèle des coefficients effectifs permet alors d’apprendre la conductivité thermique, la chaleur spécifique ainsi que le coefficient de transfert convectif à partir des données mêmes, tout en assurant la précision des prédictions.

Si l’on considère l’EDP (3.10) pour une géométrie 3-D représentant un ruban HTS uniforme, le MCE est alors capable de reproduire la solution 1-D au centre du ruban, contrairement à la méthode des différences finies. Toutefois, les réseaux de neurones du MCE $(\mathcal{C}_\theta(T), \kappa_\theta(T), \tilde{h}_\theta(T))$ n’apprennent alors plus la capacité thermique $\kappa_\theta(T)$, la chaleur spécifique $\mathcal{C}_\theta(T)$ ni le coefficient de transfert convectif $\tilde{h}_\theta(T)$ théorique, mais plutôt des coefficients effectifs, comme présenté à la figure 3.7b.

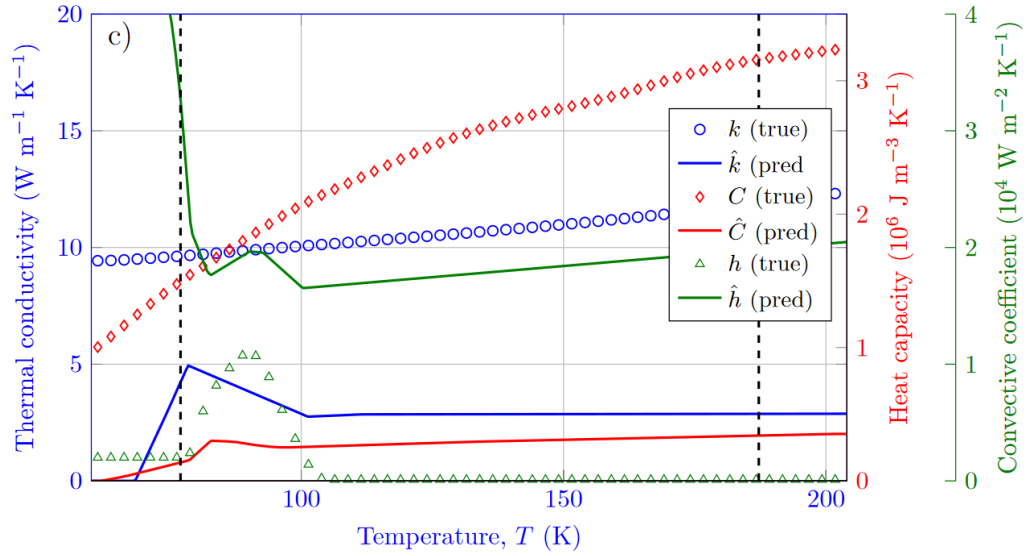
3.4 La supraconductivité par apprentissage profond

Les méthodes d’apprentissage profond offrent aujourd’hui la possibilité d’étudier les phénomènes difficilement modélisables ou inexpliqués de la supraconductivité à partir de données recueillies. En ce sens, plusieurs études cherchent à exprimer à l’aide de l’apprentissage profond la relation entre la température critique d’un HTS et son espace chimique [45, 46], tandis que d’autres tentent plutôt de prédire la force de lévitation magnétique de HTS pour des applications de Maglev [47, 48]. Certains ouvrages ont également développé des modèles

d'AP qui détermine la probabilité d'occurrence d'un quench lorsqu'un ruban est soumis à des conditions d'opérations connues [49–51]. Comme les modèles à base d'éléments finis sont souvent trop lents pour des applications en temps réel, les modèles par AP deviennent alors un outil redoutable. Toutefois, il ne semble pas exister à ce jour de modèle d'apprentissage profond qui décrit la dynamique du phénomène de quench. Comme les méthodes traditionnelles éprouvent souvent elles-mêmes de la difficulté à résoudre cette dynamique, les approches d'AP semblent a priori peu propices à surpasser celles-ci en termes de rapidité et de précision des solutions obtenues. Cependant, les chapitres suivants démontrent le contraire : il est possible de concevoir un modèle à base d'apprentissage profond en mesure de reproduire de façon précise la dynamique du quench dans un temps de calcul de l'ordre des secondes.



(a) Coefficients thermiques appris par le MCE pour une géométrie 1-D.



(b) Coefficients thermiques appris par le MCE pour une géométrie 3-D d'un ruban HTS uniforme.

FIGURE 3.7 Propriétés thermiques apprises par les réseaux de neurones pour des problèmes 1-D et 3-D. Les coefficients k , C et h désignent la conductivité thermique, la chaleur spécifique et le coefficient de transfert convectif respectivement. Figure tirée de [44].

CHAPITRE 4 MODÉLISATION DE LA DYNAMIQUE DE QUENCH PAR L'APPRENTISSAGE PROFOND

Ce quatrième chapitre introduit l'architecture du modèle proposé afin de modéliser la dynamique électrothermique de quench pour une fraction du coût de calcul d'une méthode par éléments finis classique. La stratégie pour accélérer les calculs est présentée en début de chapitre. Par la suite, le problème par éléments finis ainsi que le modèle de référence sont posés. Le chapitre se conclut avec la présentation du modèle d'apprentissage profond guidé par la physique développée et sa propriété d'extension du domaine.

4.1 Stratégies d'accélération des calculs

L'enjeu majeur avec la modélisation de quench par éléments finis est que, bien que représentative des phénomènes expérimentaux, cette méthode requiert plusieurs heures de calcul pour obtenir une seule simulation, si celle-ci converge. Le comportement résistif à faible courant qui mène à une matrice mal conditionnée et la physique multi-échelle sont deux des principales raisons de ces difficultés numériques. Par conséquent, lors de la conception de rubans HTS, où l'on étudie un espace varié de paramètres géométriques, la MEF classique est peu performante. Dans le but de remédier à ce problème, deux stratégies numériques ont été retenues afin de concevoir un modèle capable d'obtenir des résultats équivalents à la MEF en une fraction du temps de calcul :

1. Remplacer les méthodes traditionnelles par les méthodes d'apprentissage profond,
2. Réduire la dimensionnalité du problème.

4.1.1 Remplacer les méthodes traditionnelles par les méthodes d'apprentissage profond

Les premières itérations du projet ont démontré que les méthodes d'AP guidées par la physique telles que les PINN ou les DeepONet ne peuvent pas reproduire précisément la non-linéarité de la dynamique de quench à elles-seules. Le comportement exponentiel de la résistivité de la couche HTS est particulièrement difficile à apprendre pour ces méthodes. En ce sens, si l'on rajoute au problème les nombreuses interactions entre les matériaux aux propriétés thermiques et électriques non linéaires en fonction de la température, ainsi que la nature multi-échelle des rubans HTS, on complexifie énormément la tâche d'apprentissage. Notons par ailleurs que, comme la physique modélisée est un emballement thermique, les modèles

utilisés sont très susceptibles à l'accumulation d'erreur. Les modèles d'AP comme ceux mentionnés sont alors à éviter puisqu'ils commettent déjà des erreurs relatives trop importantes à partir des premières itérations. Une méthode reposant uniquement sur de l'apprentissage profond semble donc à exclure.

4.1.2 Réduire la dimensionnalité du problème

La deuxième stratégie consiste quant à elle à réduire la dimensionnalité du problème étudié afin de diminuer de plusieurs ordres de grandeur la taille matricielle des systèmes à résoudre. Comme la géométrie des rubans supraconducteurs est quasi-planaire, voire quasi-linéique, on s'intéresse davantage aux solutions longitudinales $T(x, t)$ et $V(x, t)$ qu'aux solutions complètes $T(x, y, z, t)$ et $V(x, y, z, t)$. Ceci est d'autant plus vrai lorsqu'on considère des rubans HTS dits « uniformes », dans lesquels les solutions sont pratiquement invariantes sur la largeur, i.e., $T(x, y, z, t) \approx T(x, z, t)$. Cette invariance n'est toutefois pas observée selon l'épaisseur tel qu'illustré à la figure 4.1. Ceci dit, bien que la réduction de dimensionnalité semble être une stratégie qui s'exprime naturellement, un modèle basé uniquement sur la réduction de dimensionnalité est voué à être inexact, comme mentionné dans la section 2.4.2.

4.1.3 La stratégie retenue

L'application individuelle des deux stratégies discutées précédemment semble inadaptée à la résolution de la dynamique de quench. Pour arriver à de meilleures performances, le nouveau modèle numérique proposé est formulé en tenant compte de ces deux stratégies simultanément.

L'objectif du modèle développé est d'appliquer une réduction de dimensionnalité afin de modéliser la solution longitudinale $T(x, t)$ et $V(x, t)$ au centre de la couche supraconductrice, tel qu'illustré à la figure 4.2. On modélise le milieu de la couche HTS puisqu'il s'agit de l'emplacement où la nucléation du quench s'initie. Toutefois, une réduction de dimensionnalité ne permet de représenter exactement le milieu d'une couche en raison de l'hypothèse d'homogénéisation selon l'épaisseur. Il est alors nécessaire d'incorporer des composantes d'apprentissage profond au modèle afin d'introduire la notion de localité. Ces composantes permettent ainsi d'apprendre la physique non modélisée à une échelle locale ainsi que de corriger les erreurs associées à la réduction de dimensionnalité.

Le modèle numérique proposé est donc un modèle hybride. Ce dernier utilise la combinaison (i) d'un modèle 1-D pour la dynamique de quench, afin d'insérer des connaissances sur la physique modélisée et (ii) de composantes d'apprentissage profond dans le but de paramétrer

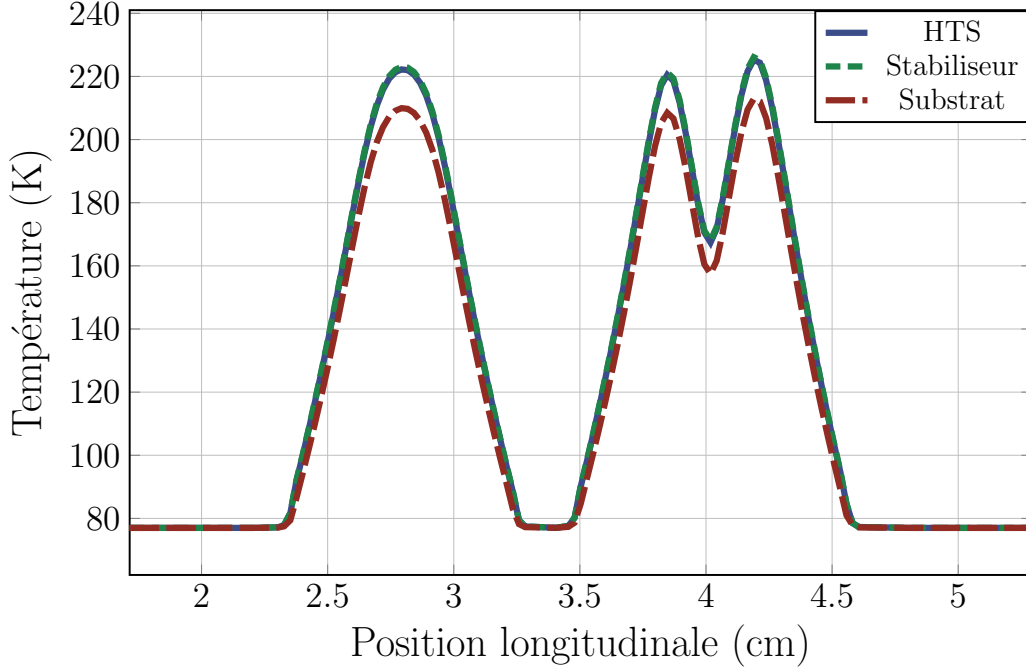


FIGURE 4.1 Température au centre des trois principales couches d'un ruban supraconducteur uniforme, soit la couche supraconductrice (HTS), le stabilisateur et le substrat lors de l'occurrence d'un quench. Les résultats ont été obtenus avec une simulation éléments finis 2-D sur un ruban comportant 3 défauts selon la longueur.

ce dernier. Notons par ailleurs que plus le modèle 1-D est initialement précis, plus le processus d'apprentissage est simplifié. Le modèle d'AP résultant de cette combinaison peut alors être entraîné à partir de simulations (2-D ou 3-D) résolues par la MEF. Ces simulations sont également séparées en deux ensembles distincts : un ensemble d'entraînement et un ensemble de test. L'ensemble d'entraînement sert de jeu de données pour le processus d'apprentissage, tandis que l'ensemble de test permet de valider la généralisation du modèle face à des données inconnues. Ce modèle hybride permet d'obtenir les solutions longitudinales $T(x, t)$ et $V(x, t)$ à très faible coût de calcul par rapport à la méthode des éléments finis. Cette dernière requiert de résoudre l'entièreté du domaine pour seulement extraire la solution au centre de la couche HTS. La figure 4.3 schématise l'idée générale de l'approche suggérée.

4.2 Modélisation de la dynamique de quench par éléments finis

Dans le cadre de ce mémoire, la dynamique de quench sera modélisée par une formulation électrothermique, i.e. un problème électrique d'écoulement de courant couplé à un problème thermique. Les sous-sections suivantes présentent le modèle d'éléments finis utilisé pour gé-

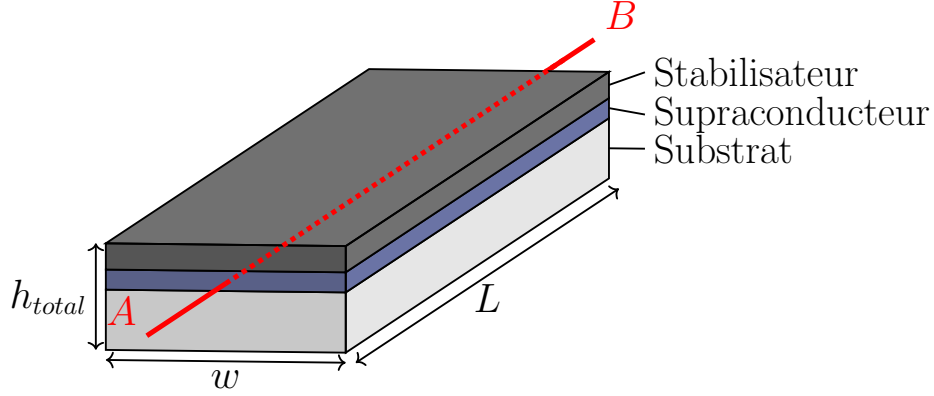


FIGURE 4.2 Schéma 3-D d'un ruban HTS uniforme : le modèle proposé apprend la solution le long de la ligne $A - B$ au centre de la couche HTS. Les variables h_{total} , w et L sont de l'ordre des μm , mm et m respectivement.

nérer les données d'entraînement ainsi que l'espace de paramètres étudiés dans le contexte d'apprentissage.

4.2.1 Simulations par éléments finis

Le modèle électrothermique par éléments finis est formulé à partir des modules *Electric currents* et *Heat transfert in solids* du logiciel COMSOL Multiphysics 6.2. Ce modèle résout les variables couplées de potentiel électrique $V(x, y, z, t)$ ainsi que de champ de température $T(x, y, z, t)$. La géométrie étudiée pour les rubans uniformes consiste en empilement structuré de couches minces formées d'un substrat, de couches tampons, d'une couche supraconductrice, d'une résistance d'interface ainsi que d'une couche de stabilisation, tel qu'illustrée à la figure 4.4. Les équations qui régissent la partie électrique du problème s'expriment ainsi :

$$\nabla \cdot (\sigma(T) \nabla V) = 0 \quad \text{dans } \Omega, \quad (4.1)$$

$$\int_{\partial\Omega_1} -\mathbf{n} \cdot (-\sigma(T) \nabla V) dS = I_{app}(t) \quad \text{sur } \partial\Omega_1, \quad (4.2)$$

$$V = 0 \quad \text{sur } \partial\Omega_2, \quad (4.3)$$

$$\mathbf{n} \cdot \nabla V = 0 \quad \text{ailleurs}, \quad (4.4)$$

où $\sigma = 1/\rho$ est la conductivité électrique, Ω est le domaine, $\partial\Omega_1$ et $\partial\Omega_2$ sont les surfaces aux extrémités du domaine, $\mathbf{n}$ est le vecteur unitaire normal à la surface considérée et $I_{app}(t)$ est

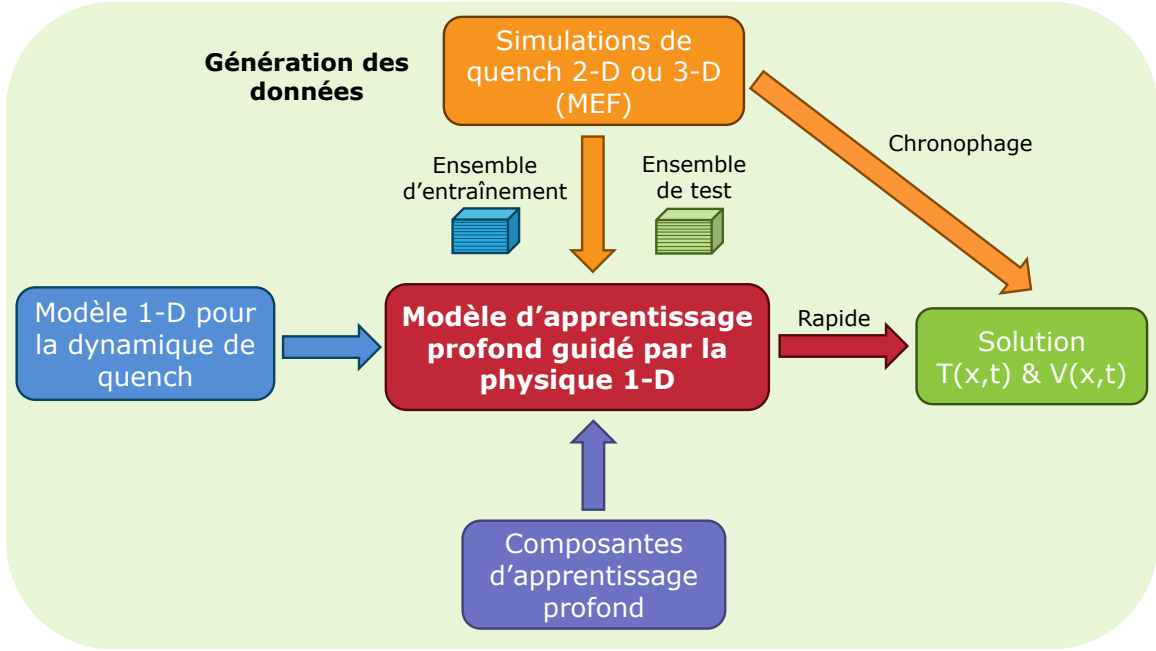


FIGURE 4.3 Schéma de la stratégie suggérée afin de remplacer la MEF par un modèle d'apprentissage profond.

le courant injecté dans le ruban. Notons que l'équation (4.1) représente la conservation du courant dans le ruban, l'équation (4.2) modélise la source de courant appliqué à l'extrémité gauche du ruban, l'équation (4.3) correspond à la mise à la terre de l'extrémité opposée, et l'équation (4.4) est une condition de Neumann qui signifie que le ruban est isolé électriquement sur toute sa longueur.

Les équations qui régissent la partie thermique du problème sont quant à elles formulées par

$$\rho_m C_m(T) \frac{\partial T}{\partial t} + \nabla \cdot (-\kappa(T) \nabla T) = \sigma(T) (\nabla V)^2 \quad \text{dans } \Omega, \quad (4.5)$$

$$T = T_0 \quad \text{sur } \partial\Omega_1 \text{ et } \partial\Omega_2, \quad (4.6)$$

$$\mathbf{n} \cdot (\kappa(T) \nabla T) = \tilde{h} \cdot (T - T_0) \quad \text{ailleurs}, \quad (4.7)$$

où ρ_m est la densité massique, C_m est la capacité thermique massique, et le terme source $\sigma(T)(\nabla V)^2$ représente la chaleur dissipée par effet Joule. L'équation (4.5) décrit l'évolution de la distribution de température T , l'équation (4.6) est une condition de Dirichlet aux extrémités du ruban (conditions frontières isothermes), et l'équation (4.7) modélise le refroidissement par convection entre les surfaces extérieures du ruban et l'environnement cryogénique maintenu à température T_0 .

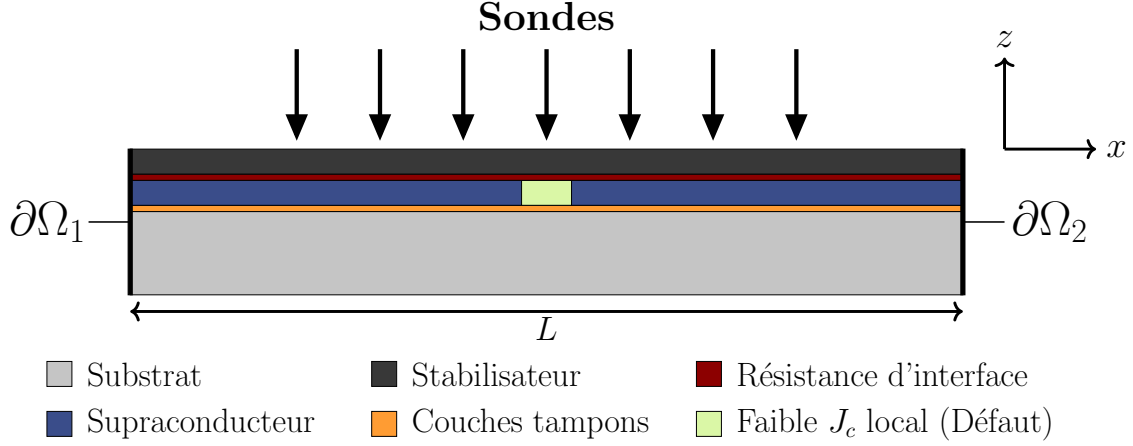


FIGURE 4.4 Schéma d'une géométrie 2-D d'un ruban HTS uniforme avec une densité de courant critique J_c faible en son centre. Les frontières du domaine (extrémités du ruban) sont dénotées par $\partial\Omega_1$ et $\partial\Omega_2$.

Les matériaux utilisés dans le modèle sont ceux qu'on retrouve régulièrement dans les rubans supraconducteurs industriels, soit l'Hastelloy en guise de substrat, le MgO pour les couches tampons, un REBCO en tant que supraconducteur de type II et l'argent pour la couche stabilisatrice. Leurs propriétés thermiques et électriques respectives ont été tirées de [12, 52–56], et celles-ci sont présentées à la figure 4.5. La conductivité électrique de la couche supraconductrice σ_{rebcO} est quant à elle modélisée par deux conductivités en parallèle telle que

$$\sigma_{rebcO}(x, T, E) = \sigma_{sc}(x, T, E) + \sigma_n, \quad (4.8)$$

où σ_{sc} est la conductivité du REBCO dans le régime transitoire tandis que σ_n est la conductivité à l'état normal. La conductivité σ_{sc} s'exprime comme une loi de puissance de la forme

$$\sigma_{sc}(x, T, E) = \frac{J_c(x, T)}{\|\mathbf{E}\|_2} \left(\frac{\|\mathbf{E}\|_2}{E_c} \right)^{\frac{1}{n(T)}}, \quad (4.9)$$

où

$$J_c(x, T) = \begin{cases} J_{c0} \frac{T_c - T}{T_c - T_0} \Phi(x) & \text{si } T < T_c \\ 0 & \text{si } T \geq T_c \end{cases}, \quad (4.10)$$

et

$$n(T) = \begin{cases} n_0 & \text{si } T < T_c \\ 1 & \text{si } T \geq T_c \end{cases}. \quad (4.11)$$

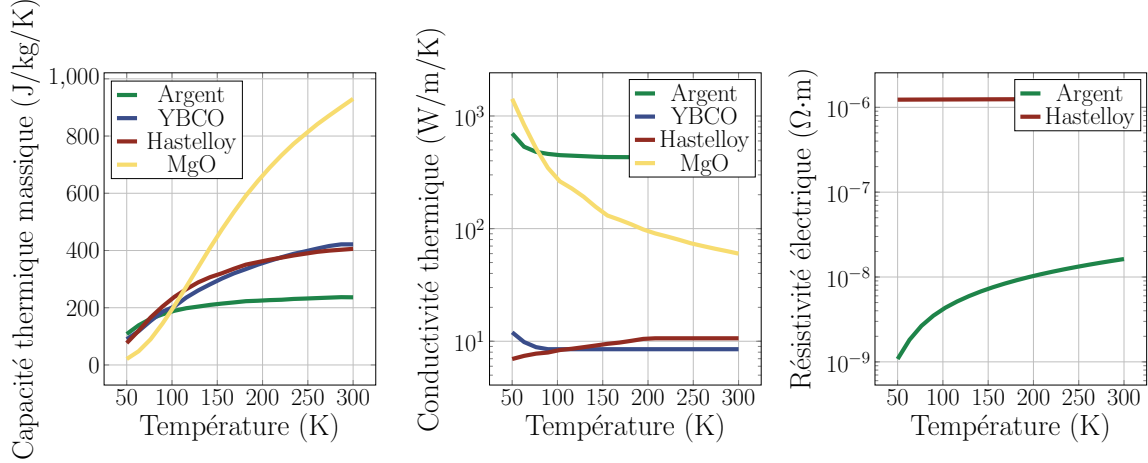


FIGURE 4.5 Propriétés thermiques et électriques utilisées dans le modèle électrothermique par éléments finis [12, 52–56]. La résistivité électrique du MgO n’est pas requise lors de l’approximation en couche mince dans COMSOL.

Notons que J_{c0} représente la densité de courant critique à $T = T_0$, n_0 est un paramètre d’ajustement propre au matériau supraconducteur et $\Phi(x)$ est une fonction qui décrit la distribution spatiale des inhomogénéités soumise à la condition $0 \leq \Phi(x) \leq 1$. Le terme $\Phi(x)$ établit des régions locales où la densité de courant critique J_c est affaiblie afin d’initier un emballement thermique, e.i. un quench lors des simulations.

Le problème électrothermique ainsi formulé est hautement non linéaire et couplé en raison des équations (4.1) et (4.5). Ceci dit, la plus forte non linéarité du problème réside dans la conductivité électrique du REBCO, qui dépend à la fois des variables de résolution V et T . Cette formulation a pour conséquence de provoquer d’abruptes partages de courant et d’importantes générations de chaleur localisés autour des zones à faibles J_c .

4.2.2 Génération des données d’entraînement

Lors de la génération de données, il est essentiel de diversifier au maximum les données d’entraînement afin de faciliter le processus d’apprentissage et la robustesse du modèle. Dans le cadre de l’étude de la dynamique de quench, cette diversification des données est obtenue en variant aléatoirement la résistance d’interface, l’épaisseur de la couche stabilisatrice, les inhomogénéités de la densité de courant critique ainsi que les rampes de courant $I_{app}(t)$. Ceci dit, les plages de variation des paramètres ont été établies empiriquement, mais notons qu’il est possible de s’appuyer sur des approches mathématiques plus rigoureuses afin de mieux couvrir l’espace de paramètres. Les paramètres géométriques des rubans considérés lors des

simulations sont présentées au tableau 4.1. La dynamique de quench est modélisée sur une durée de 10 ms, puisque c'est à cette échelle de temps que les phénomènes de nucléation de quench et de propagation de la zone normale sont généralement observés dans les rubans HTS possédant une mince couche de stabilisateur.

TABLEAU 4.1 Paramètres géométriques de rubans HTS uniformes.

Paramètres	Valeur numérique
Longueur (L)	7 cm
Largeur (w)	4 mm
Épaisseur de substrat (h_{hast})	50 μm
Épaisseur couche tampon (h_{MgO})	150 nm
Épaisseur couche supraconductrice (h_s)	2 μm
Épaisseur résistance d'interface (h_{int})	[100,1000] nm
Épaisseur couche stabilisatrice (h_a)	[1.65,2.35] μm

Paramètres géométriques aléatoires

La plage de variation de l'épaisseur de la résistance d'interface présentée au tableau 4.1 correspond à une résistance d'interface variant entre 1 et 10 $\mu\Omega.\text{cm}^2$. Le choix de cet intervalle garantit que la CTL engendrée par R_{inter} a un impact notable sur une longueur de 7 cm, tout en assurant la présence de zones tampons aux extrémités du domaine. Ces zones tampons empêchent l'apparition d'effets de bord pouvant potentiellement fausser la physique simulée. Par ailleurs, la longueur du ruban est limitée à 7 cm afin que la durée de chaque simulation demeure raisonnable ($\sim$ 1-3 heures).

La variation de l'épaisseur de la couche d'argent est une autre façon de faire varier les profils de température observés lors d'un quench par le biais de l'architecture du ruban. Si l'épaisseur h_a est faible, alors la densité de courant à l'intérieur du stabilisateur J_a est élevée lors de l'occurrence d'un quench, ce qui résulte en une génération de chaleur locale importante. Toutefois, si l'épaisseur est grande, l'ajout de masse thermique et la réduction de la résistance électrique entraînent une densité de courant J_a plus faible, une génération de chaleur modeste, et une élévation de température moindre. Notons que la plage de valeurs de h_a a été déterminée de façon à ce que les valeurs extrêmes engendre au maximum une élévation de température de 150 K sur la température maximale simulée.

Inhomogénéités aléatoires

Les inhomogénéités de la densité de courant critique J_c dictent l'emplacement ainsi que la taille des zones résistives où le partage de courant a lieu lors d'un quench, régissant alors la distribution de température T dans les diverses couches. En ce sens, la largeur (σ_i), l'amplitude ($\bar{A}_i$), la position longitudinale ($x_{c,i}$) ainsi que le nombre (n_D) d'inhomogénéités (aussi appelées défauts) sont variés aléatoirement afin de pouvoir représenter la plus grande famille de configurations de J_c possible. Pour ce faire, la distribution spatiale des inhomogénéités du REBCO est modélisée par la combinaison de n_D inhomogénéités gaussiennes telles que

$$\Phi(x) = \prod_{i=1}^{n_D} \left\{ 1 - \bar{A}_i \exp \left[- \frac{(x - x_{c,i})^2}{2\sigma_i^2} \right] \right\}. \quad (4.12)$$

Toutes les valeurs numériques utilisées pour la couche supraconductrice du REBCO sont présentées dans le tableau 4.2. On remarque que le profil $\Phi(x)$ peut devenir rapidement complexe en fonction du nombre de défauts, tel qu'illustré à la figure 4.6. Par ailleurs, notons qu'il est tout à fait possible que deux ou plusieurs zones normales en cours de propagation originant de défauts différents se rejoignent et deviennent une seule grande zone normale lors d'un emballement thermique sévère ou prolongé. Dans ce cas, la taille des défauts résultants peut être plus grande que la taille maximale indiquée dans le tableau 4.2. Notons également que la plage de variation de l'amplitude des défauts $\bar{A}_i$ a été déterminée de façon à ce que la température s'élève de T_0 à environ 300 K sur une échelle de 10 ms, tout en conservant des défauts de taille comparable. De cette façon, on évite la présence de défauts importants ($\bar{A}_i = 1$) pouvant rapidement englober les défauts de tailles moindres ($\bar{A}_i < 0.5$), ce qui compromettrait la variabilité du jeu de données.

TABLEAU 4.2 Paramètres du REBCO.

Paramètres	Valeur numérique
Température critique (T_c)	90 K
Température de l'environnement cryogénique (T_0)	77 K
Coefficient de transfert thermique ($\tilde{h}$)	2000 W m ⁻²
Densité de courant critique à T_0 (J_{c0})	2.5 MA cm ⁻²
Paramètre d'ajustement (n_0)	20
Largeur du défaut i (σ_i)	[0.7,3.2] mm
Amplitude du défaut i ($\bar{A}_i$)	[0.55,0.75]
Position du défaut i ($x_{c,i}$)	[0.4,0.6]· L cm
Nombre de défauts (n_D)	2,3,4

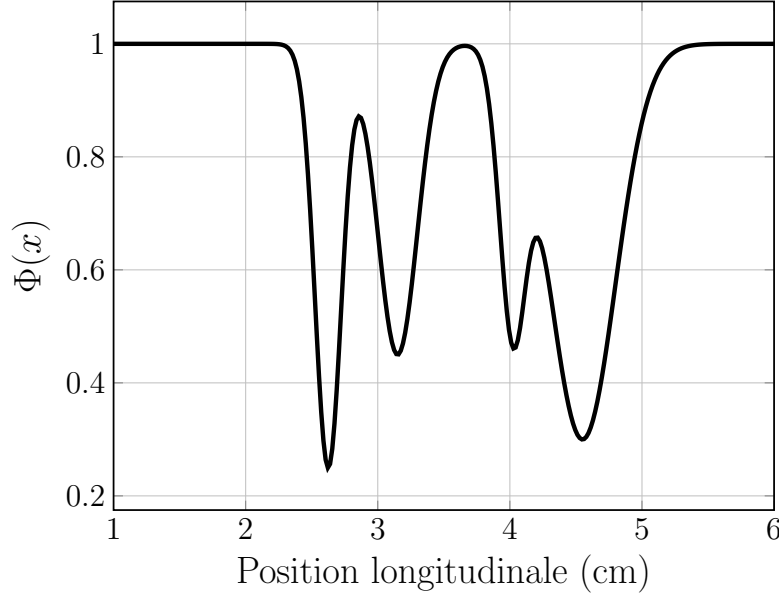


FIGURE 4.6 Exemple d'un profil d'inhomogénéités $\Phi(x)$ pour $n_D = 4$ défauts gaussiens.

Rampes de courant aléatoires

Les rampes de courant $I_{app}(t)$ appliquées aux rubans HTS sont directement corrélées à la température maximale qu'un quench peut atteindre. Il est ainsi pertinent de s'assurer de faire varier ces rampes si l'on désire diversifier les données d'entraînement. Ceci dit, les travaux préliminaires de ce mémoire ont démontré qu'il est préférable d'avoir des rampes de courant qui varient dans le temps ($\frac{dI_{app}}{dt} \neq 0$) plutôt que des rampes constantes lors de l'entraînement. Les rampes constantes ont pour conséquence d'induire une dissipation de chaleur par effet Joule relativement constante, réduisant alors la variabilité du jeu de données.

Les rampes de courant déterminent également le régime de stabilité thermique dans lequel le ruban se retrouve, tel qu'expliqué dans la section 2.2.4. Pour cette raison, les rampes seront conçues de sorte à modéliser deux régimes : le régime de *quench* et le régime de refroidissement (*cryostabilité* et *cold-end recovery*). Le régime de quench correspond au scénario où les pertes ohmiques sont dominantes, tandis que le régime de refroidissement modélise le scénario où le refroidissement par convection est localement équivalent ou plus élevé que les pertes ohmiques. Ces deux scénarios sont illustrés à la figure 4.7.

L'espace de rampes utilisées pour la génération de données est présenté à la figure 4.8. Chaque

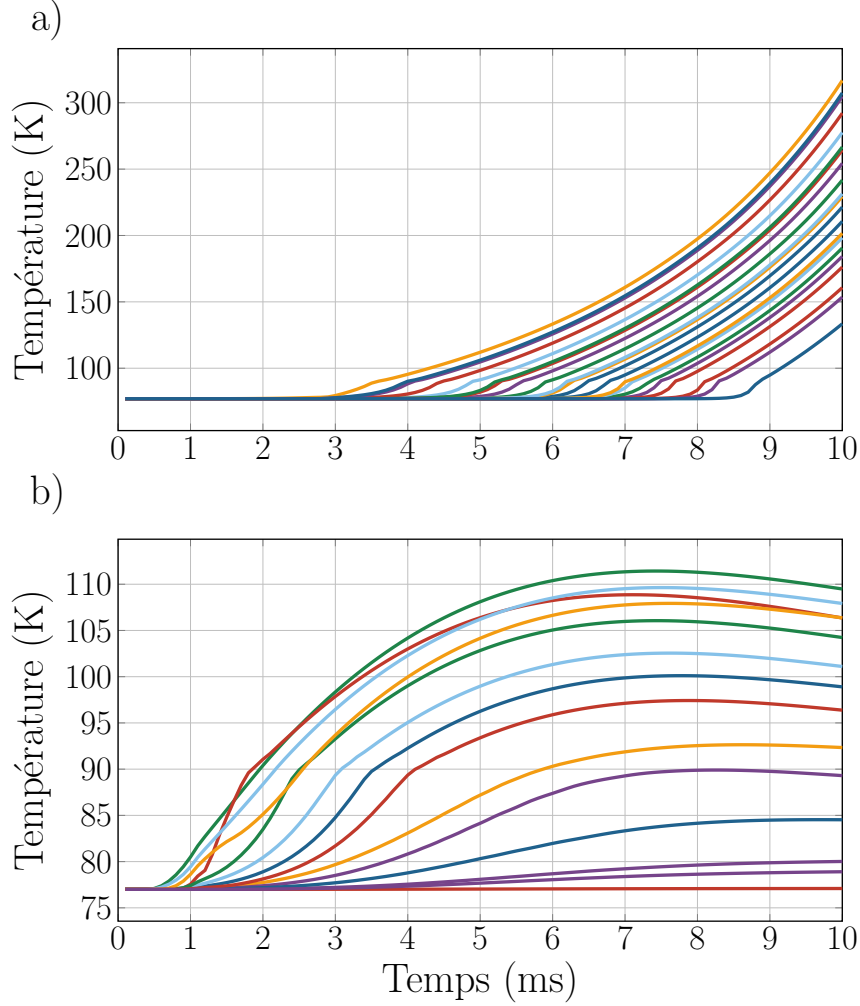


FIGURE 4.7 Évolution de la température en fonction du temps pour les régimes (a) dominés par les pertes Joule et (b) dominés par le refroidissement. Chacune des courbes correspond à la température lue par les sondes distribuées uniformément le long du ruban.

rampe de courant est modélisée par une fonction par partie telle que

$$I_{app}(t) = \begin{cases} m_1 t + b_1 & \text{si } 0 \leq t < 1 \text{ ms} \\ m_2 t + b_2 & \text{si } 1 \leq t \leq 10 \text{ ms} \end{cases}. \quad (4.13)$$

où m est la pente de la droite, b est l'ordonnée à l'origine, et où l'on impose la continuité de $I_{app}(t)$. La première droite, à pente abrupte, sert à initialiser un emballement thermique tandis que la seconde permet de déterminer la nature du régime (quench ou refroidissement). Les rampes dominées par les pertes Joule sont conçues de manière à ce que la température maximale après 10 ms soit comprise en moyenne entre 175 K et 350 K. Ces rampes sont ca-

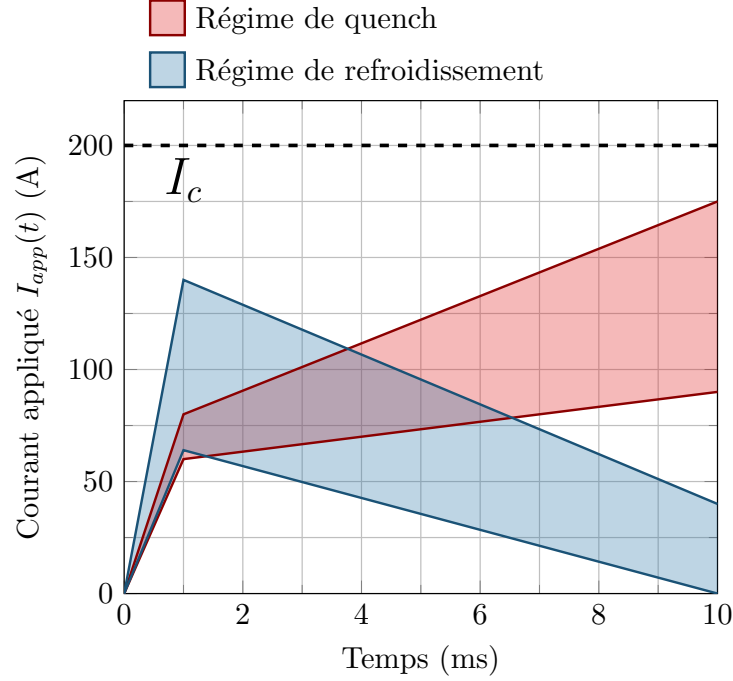


FIGURE 4.8 Intervalles des rampes de courant décrivant un régime de quench (rouge) et un régime de refroidissement (bleu).

ractérisées par une température strictement croissante. Pour les rampes où le refroidissement devient dominant, la température maximale au dernier pas de temps est comprise entre T_0 et 160 K. L'évolution de la température n'est alors pas nécessairement monotone croissante. Notons que le modèle hybride considéré extrapole si la rampe de courant $I_{app}(t)$ utilisée est en dehors de l'espace défini à la figure 4.8. Ceci dit, les données sont assez diversifiées afin de permettre de construire un modèle robuste, capable d'extrapoler de façon raisonnable.

4.3 Modèle de référence

Un modèle de référence permet de définir les performances minimales qu'un modèle d'apprentissage profond doit surpasser afin d'être considéré comme une amélioration. Dans le cadre de ce mémoire, le modèle 1-D de Levin *et al.* présenté à la section 2.4.1 sera le modèle de référence utilisé. Cette section a pour objectif de présenter les détails d'implémentation pertinents du modèle de référence ainsi que la raison de ces choix.

D'une part, l'équation thermique (2.3) du modèle électrothermique 1-D est implémentée à

partir de la discrétisation de Du Fort-Frankel, telle que

$$T_i^{n+1} = \frac{1}{1 + A_i} \left[(1 - A_i) T_i^{n-1} + A_i (T_{i+1}^n + T_{i-1}^n) + B_i (T_{i+1}^n - T_{i-1}^n)^2 + \gamma_i Q_i^n \right], \quad (4.14)$$

où

$$A_i = \frac{2K(T_i^n) \Delta t}{C(T_i^n) \Delta x^2}, \quad B_i = \frac{\Delta t}{2C(T_i^n) \Delta x^2} \frac{\partial K(T_i^n)}{\partial T}, \quad \gamma_i = \frac{2\Delta t}{C(T_i^n)}, \quad (4.15)$$

et

$$Q_i^n = Q_{a,i}^n + Q_{s,i}^n + Q_{inter,i}^n + Q_{conv,i}^n. \quad (4.16)$$

Le développement détaillé se retrouve à l'annexe A. On privilégie la méthode de Du Fort-Frankel à celle de Crank-Nicolson en raison de la nature explicite de la discrétisation. Une discrétisation explicite se prête mieux à d'éventuels couplages entre les algorithmes d'apprentissage profond et les différences finies en raison de leur faible coût de calcul, similaire à celui d'un réseau de neurones *feed forward*. De leur côté, les méthodes implicites requièrent de résoudre un système matriciel pouvant devenir chronophage lorsque l'on considère de longs rubans. De plus, on remarque que l'équation (4.14) nécessite de connaître les solutions $T(x, 0)$ et $T(x, \Delta t)$ préalablement à la résolution du problème. Il ne s'agit toutefois pas d'un enjeu car ces solutions sont toutes deux triviales (aucune génération de chaleur car $J_s < J_c \forall x \in \Omega$) pour les premiers instants.

D'autre part, l'équation électrique transcendante (2.4) peut être résolue de façon itérative par un algorithme de point fixe (convergence d'ordre 1) ou de Newton-Raphson (convergence à l'ordre 2). Des résultats préliminaires ont toutefois démontré que l'algorithme de point fixe converge plus difficilement en raison du comportement exponentiel de ρ_s , et donc l'algorithme de Newton-Raphson a été priorisé. On cherche donc à résoudre l'EDP formée des équations (2.4) et (2.6) à chaque instant t telles que

$$r(J_s) = \underbrace{-R_{inter} h_a \frac{A_s}{A_a} \frac{\partial^2 J_s}{\partial x^2}}_{\eta} - \underbrace{(\rho_a J_a - \rho_s(J_s) J_s)}_{\mathcal{S}(J_s)} = 0, \quad (4.17)$$

où A_a et A_s sont les sections transversales des couches stabilisatrice et supraconductrice respectivement, et on suppose que R_{inter} , A_a et A_s sont invariantes selon la longueur. En approximant la dérivée seconde de J_s par des différences finies centrées, on obtient un système de $N_x - 2$ équations non linéaire telle que

$$r_i = \eta \frac{J_{s,i-1} - 2J_{s,i} + J_{s,i+1}}{\Delta x^2} - \mathcal{S}(J_{s,i}) = 0 \quad \text{pour } i = 2, 3, \dots, N_x - 1, \quad (4.18)$$

où N_x est le nombre de points de maillage. Notons par ailleurs que deux conditions de Dirichlet sont imposées aux extrémités du domaine. La résolution de l'équation (4.18) à l'aide de la méthode de Newton-Raphson consiste alors à résoudre le système matriciel

$$\mathcal{J}\Delta\mathbf{J} = -\mathbf{r} \quad (4.19)$$

à chaque itération k de l'algorithme, où $\mathcal{J}$ est la matrice Jacobienne du système à résoudre, $\Delta\mathbf{J}$ est le vecteur de correction inconnu, et $\mathbf{r}$ est le vecteur résidu associé aux équations r_i . Il est également aisé de démontrer que la matrice $\mathcal{J}$ est tridiagonale et s'exprime comme

$$\mathcal{J}_{ij} = \begin{cases} -\frac{2\eta}{\Delta x^2} - \frac{\partial \mathcal{S}}{\partial J_s} & \text{si } i = j \\ \frac{\eta}{\Delta x^2} & \text{si } i = |j \pm 1|, \\ 0 & \text{sinon} \end{cases} \quad (4.20)$$

et que

$$\frac{\partial \mathcal{S}}{\partial J_s} = -\frac{A_s}{A_a}\rho_a - n\rho_s. \quad (4.21)$$

L'équation (4.19) est résolue à l'aide de l'algorithme de Thomas, une forme simplifiée de l'élimination de Gauss pour les systèmes tridiagonaux en $\mathcal{O}(n)$. La solution discrétisée $\mathbf{J}_s$ à l'itération $k + 1$ est alors simplement

$$\mathbf{J}_s^{(k+1)} = \mathbf{J}_s^{(k)} + \alpha_{NR}\Delta\mathbf{J}, \quad (4.22)$$

où α_{NR} est un facteur d'amortissement. L'algorithme complet est présenté à l'annexe B. Notons également que le critère d'arrêt utilisé est l'erreur relative dans la norme L_2 entre deux itérations consécutives.

Un schéma de la résolution séquentielle découplée du modèle de référence est présenté à la figure 4.9. Le modèle de référence sera également référé à la méthode des différences finies 1-D, ou MDF 1-D, en raison de la nature de son implémentation. Par ailleurs, on remarque que le temps nécessaire à la résolution du modèle de référence est majoritairement dominé par celui associé à la résolution de l'équation (4.19). Ceci est expliqué par la nature implicite de l'algorithme de Newton-Raphson.

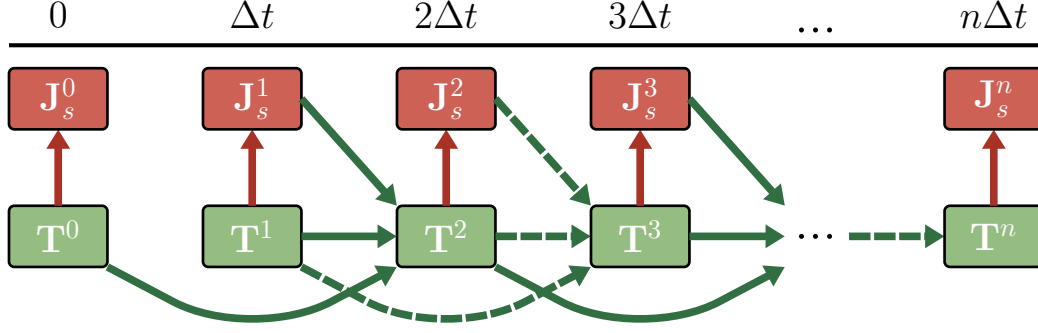


FIGURE 4.9 Schéma de la résolution séquentielle dé耦ée du modèle de référence. Les flèches vertes et rouges réfèrent aux équations (4.14) et (4.18), respectivement. Les lignes pointillées servent seulement à différencier les instants.

4.4 Modèle d'apprentissage profond guidé par la physique pour la dynamique de quench

Le modèle d'apprentissage profond guidé par la physique proposé consiste à paramétrer uniquement le problème thermique du modèle de référence. Le problème électrique ne se prête pas à une paramétrisation par réseaux de neurones en raison de sa nature implicite et transcendante. Par ailleurs, paramétrer autant le problème thermique qu'électrique ne ferait qu'alourdir les calculs et inutilement complexifier le modèle. Si on suppose que le problème thermique paramétré permet d'obtenir la solution exacte $T(x, t)$, alors la résistivité électrique $\rho(T)$ est également connue. La solution $J_s(x, t)$ peut alors être reconstruite par l'équation (4.18) jusqu'à une tolérance désirée, et le problème électrique ne nécessite donc pas d'être paramétré.

En ce sens, le modèle hybride proposé sera appelé modèle « **R**écurrent d'**A**pprentissage profond guidé par la **P**hysique basé sur les **D**ifférences finies » (RAPDi). Plus précisément, le modèle de Levin *et al.* sert à « guider la solution » à l'aide de contraintes physiques dures appliquées sous la forme de différences finies, alors que des composantes d'apprentissage profond sont incorporées au sein de ce même modèle (équation (4.14)) afin de le paramétrer. Également, comme les réseaux de neurones récurrents peuvent être vus comme un système dynamique à temps discret, la résolution séquentielle dé耦ée (figure 4.9) peut être interprétée comme un RNN, tel qu'illustré à la figure 4.10 (côté droit). On remarque qu'il existe ainsi une analogie entre la solution passée $\mathbf{T}^{n-1}$ et les états cachés $\mathbf{h}_t$ d'un RNN. D'un point de vue de l'apprentissage profond, on peut dire que ces deux quantités établissent un « contexte » sur la dynamique étudiée.

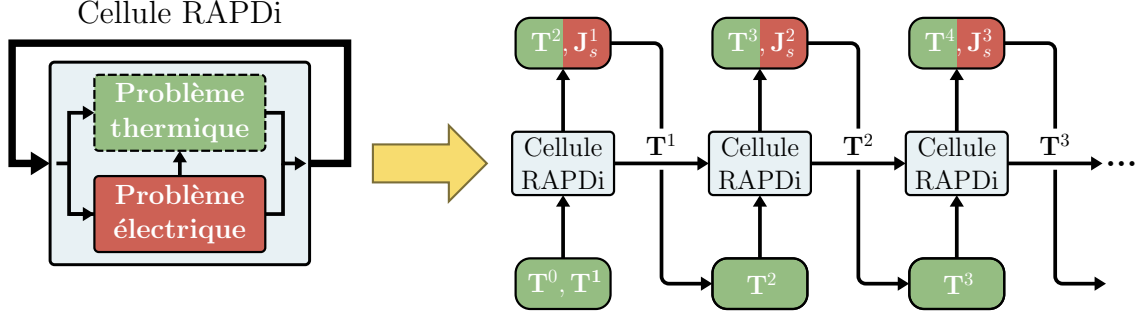


FIGURE 4.10 Schéma de l'intérieur d'une cellule RAPDi avec connexion récurrente à gauche et le dépliage temporel du modèle RAPDi à droite. Les lignes pointillées indiquent que des composantes d'AP sont incorporées.

Les prochaines sous-sections présentent les inconvénients du modèle des coefficients effectifs ainsi que la reformulation du modèle des coefficients effectifs (MCE) qui a inspiré le modèle RAPDi. Par la suite, le modèle RAPDi et sa propriété d'extension du domaine sont présentés, suivis des stratégies d'entraînement utilisées.

4.4.1 Inconvénients du modèle des coefficients effectifs

Le MCE présenté à la section 3.3.2 n'est pas utilisé afin de paramétrer le problème thermique pour plusieurs raisons. Tout d'abord, les premières étapes du projet ont démontré que le MCE semble incapable d'apprendre les coefficients thermiques exacts qui régissent un problème électrothermique 1-D. En comparaison, lorsqu'on considère seulement un problème thermique 1-D de la forme

$$C(T) \frac{\partial T}{\partial t} + \frac{\partial}{\partial x} \left(-K(T) \frac{\partial T}{\partial x} \right) = Q(x, t), \quad (4.23)$$

où $Q(x, t)$ est une source de chaleur gaussienne, chaque terme est convenablement approximé par le MCE. Plus précisément, l'équation (3.11) approxime les dérivées spatiales et temporelles par des différences finies, tandis que le terme source $Q(x, t)$ est décrit par des fonctions élémentaires. Les réseaux de neurones $C_\theta(T)$ et $K_\theta(T)$ du MCE n'ont alors pas le choix que de converger vers les courbes réelles $C(T)$ et $K(T)$ si l'on veut minimiser l'erreur lors de l'entraînement. Toutefois, lorsqu'on considère un problème électrothermique 1-D, la source de chaleur $Q(x, T, J_s)$ ne s'exprime plus nécessairement en termes de fonctions élémentaires. Dans une telle situation, on observe que la non linéarité de $Q(x, T, J_s)$ et l'erreur associée aux découplages des équations induisent une erreur qui complexifie le processus d'apprentissage. Les coefficients appris $(C_\theta(T), K_\theta(T))$ ne convergent alors plus vers les coefficients réels $(C(T), K(T))$. Les coefficients appris restent toutefois « effectifs ».

Par ailleurs, un deuxième inconvénient est que le modèle des coefficients effectifs ne distingue pas un régime de quench d'un régime de refroidissement lorsqu'il apprend les coefficients $C_\theta(T)$ et $K_\theta(T)$. Les propriétés thermiques apprises ne sont pas nécessairement les mêmes lorsqu'on veut décrire un régime de quench où la température augmente exponentiellement (faible $C(T)$), ou un régime de refroidissement avec une température stable ($C(T)$ plus élevé). Ce conflit survient principalement pour la plage de températures entre T_0 et 150 K et empêche de représenter simultanément ces deux régimes à partir d'un même modèle.

Finalement, le MCE a également l'inconvénient d'optimiser les quantités $C_\theta(T)$ et $K_\theta(T)$ sous forme de ratio à l'intérieur des variables A_θ , B_θ et γ_θ . Cette dépendance entre la conductivité thermique effective $K_\theta(T)$ et sa dérivée sous forme de ratio rend l'optimisation difficile et mal régularisée.

4.4.2 Reformulation du modèle des coefficients effectifs

Les observations précédentes ont conduit à reformuler le modèle des coefficients effectifs. Dans cette reformulation, on suppose qu'il existe des fonctions $c(T)$, $\alpha(T)$ et $\beta(T)$ telles que les identités

$$C_\theta(T) = \frac{C(T)}{c(T)}, \quad K_\theta(T) = \alpha(T)K(T) \quad \text{et} \quad \frac{\partial K_\theta}{\partial T} = \beta(T)\frac{\partial K}{\partial T} \quad (4.24)$$

sont vérifiées. Si de telles fonctions existent, celles-ci permettent de faire le lien entre ce nouveau modèle et le MCE. Par ailleurs, cette reformulation possède un avantage : les formes fonctionnelles des propriétés thermiques sont explicitement données, ce qui préserve mieux la physique et facilite le processus d'apprentissage. Les fonctions $c(T)$, $\alpha(T)$ et $\beta(T)$ peuvent être interprétées comme des perturbations multiplicatives des propriétés réelles. Il est alors possible d'écrire les coefficients A_θ , B_θ et γ_θ de l'équation (3.12) tels que

$$A_\theta = \frac{2K_\theta(T)\Delta t}{C_\theta(T)\Delta x^2} = \frac{2K(T)\Delta t}{C(T)\Delta x^2}\alpha(T)c(T) = A_i\alpha(T)c(T), \quad (4.25)$$

$$B_\theta = \frac{\Delta t}{2C_\theta(T)\Delta x^2} \frac{\partial K_\theta}{\partial T} = \frac{\Delta t}{2C(T)\Delta x^2} \frac{\partial K}{\partial T}\beta(T)c(T) = B_i\beta(T)c(T) \quad (4.26)$$

et

$$\gamma_\theta = \frac{2\Delta t}{C_\theta(T)} = \frac{2\Delta t}{C(T)}c(T) = \gamma_i c(T). \quad (4.27)$$

La reformulation du MCE cherche alors à approximer directement les fonctions $a(T) = \alpha(T)c(T)$, $b(T) = \beta(T)c(T)$ et $c(T)$ par des réseaux de neurones notés $a_\theta(T)$, $b_\theta(T)$ et $c_\theta(T)$,

respectivement. Les coefficients A_θ , B_θ et γ_θ sont alors exprimés tels que

$$A_\theta = A_i a_\theta(T), \quad B_\theta = B_i b_\theta(T) \quad \text{et} \quad \gamma_\theta = \gamma_i c_\theta(T). \quad (4.28)$$

Cette reformulation du MCE ne se veut ni rigoureuse ni une preuve formelle, mais simplement une reformulation qui reprend l'essentiel du MCE. On ne travaille alors plus avec des propriétés thermiques apprises, mais avec des fonctions a_θ , b_θ et c_θ . Cette approche limite beaucoup les possibilités de dégénérescences du modèle. On remarque notamment que l'équation (4.28) permet de simplifier l'optimisation en limitant les ratios formés par les variables d'intérêt (exemple : $K_\theta(T)/C_\theta(T)$).

4.4.3 Conservation de la chaleur

La reformulation proposée du modèle des coefficients effectifs a pu soulever plusieurs questions pertinentes au cours de ce mémoire quant à la conservation de l'énergie et la conservation de chaleur. Lorsqu'on modélise un phénomène physique à l'aide d'une homogénéisation des propriétés thermiques, le terme de chaleur Q est inchangé, seulement les propriétés $C(T)$ et $K(T)$ sont modifiées. En ce sens, si la chaleur était conservée sans l'homogénéisation, elle l'est tout autant avec l'homogénéisation. Cependant, qu'en est-il du MCE et de sa reformulation ? En réalité, si on s'attarde sur le terme

$$\gamma_\theta Q_i^n = \underbrace{\left(\frac{2\Delta t}{C(T)} c_\theta(T) \right)}_{\text{Matériau effectif}} Q_i^n = \frac{2\Delta t}{C(T)} \underbrace{\left(c_\theta(T) Q_i^n \right)}_{\text{Pondère la chaleur}}, \quad (4.29)$$

on remarque que la question « La chaleur est-elle conservée ? » dépend de notre école de pensée. Que l'on modélise le phénomène par un matériau effectif ou que l'on pondère la chaleur, ces deux approches sont en réalité deux faces d'une même médaille dans ce contexte. En ce sens, si l'on admet qu'utiliser un matériau effectif est acceptable au sens physique, on doit également admettre par l'équation (4.29) que pondérer la chaleur Q_i^n l'est tout autant. Ce résultat parfois contre-intuitif est la raison pour laquelle le modèle RAPDi n'assure pas formellement la conservation de la chaleur. Ceci dit, une réflexion plus approfondie serait de mise afin de mieux comprendre les conséquences d'une telle interprétation. Il est toutefois pertinent de préciser que la chaleur est bien conservée dans le problème de différence finie en raison de la formulation numérique et la physique du problème. Lorsqu'on mentionne que « la chaleur n'est pas conservée par le modèle RAPDi », on réfère au fait que les chaleurs calculées par le modèle RAPDi peuvent être différentes des chaleurs calculées par le problème éléments finis.

4.4.4 Le modèle RAPDi

Le modèle RAPDi paramétrise seulement le problème thermique du modèle de référence à partir des « préfacteurs » $a_\theta, b_\theta, c_\theta, d_\theta, e_\theta$ et f_θ tel que

$$T_{\theta,i}^{n+1} = \frac{1}{1 + A_\theta} \left[(1 - A_\theta) T_i^{n-1} + A_\theta (T_{i+1}^n + T_{i-1}^n) + B_\theta (T_{i+1}^n - T_{i-1}^n)^2 + \gamma_\theta Q_i^n \right], \quad (4.30)$$

où

$$A_\theta = A_i a_\theta, \quad B_\theta = B_i b_\theta, \quad (4.31)$$

et

$$\gamma_\theta Q_i^n = \gamma_i \left[(Q_{ag,i}^n + Q_{s,i}^n) c_\theta + Q_{inter,i}^n d_\theta + \frac{Q_{conv,i}^n}{2} e_\theta + \frac{Q_{conv,i}^n}{2} f_\theta \right]. \quad (4.32)$$

Les préfacteurs ne sont plus des fonctions apprises telles que $a_\theta(T)$ ou $b_\theta(T)$, mais des éléments du vecteur de sortie d'un modèle d'apprentissage profond. Le préfacteur a_θ (A_θ) aide en partie à la stabilité de la méthode numérique, b_θ pondère le gradient de la solution ($T_{i+1}^n - T_{i-1}^n$) et les autres préfacteurs pondèrent les différents termes de chaleur. L'intuition d'incorporer deux termes $Q_{conv,i}^n$ est que l'un des préfacteurs contrôle le régime dominé par les pertes Joule, tandis que le second contrôle le régime de refroidissement. On constate ainsi que le modèle RAPDi ne partage plus de lien avec l'interprétation d'un matériau effectif en raison de l'apparition des préfacteurs d_θ, e_θ et f_θ . Cependant, on remarque que si les préfacteurs d_θ, e_θ et f_θ convergent vers c_θ , alors le modèle RAPDi est équivalent à la reformulation du MCE. Les préfacteurs peuvent être interprétés comme des degrés de liberté qui améliorent la précision de l'équation (4.14).

Le modèle d'apprentissage profond utilisé à l'intérieur du modèle RAPDi est présenté à la figure 4.11. On introduit de ce fait-même la notation $\mathbf{T}_j^n$ qui désigne la température T au temps t_n pour une fenêtre de $j+1$ points centrée sur le point x_i du maillage. À titre d'exemple, l'expression

$$\mathbf{T}_4^n = [T_{i-2}^n, T_{i-1}^n, T_i^n, T_{i+1}^n, T_{i+2}^n]^T, \quad (4.33)$$

représente un vecteur formé de T_i^n et ses 4 plus proches voisins. Les entrées du modèle d'apprentissage profond sont donc les variables $\mathbf{T}_j^{n-1}$, $\mathbf{T}_j^n$, $\mathbf{Q}_{ag,j}^n$, $\mathbf{Q}_{s,j}^n$, $\mathbf{Q}_{inter,j}^n$ et $\mathbf{Q}_{conv,j}^n$. Celles-ci permettent de tenir compte à la fois du contexte spatial et du contexte thermique. Ces informations sont essentielles afin de distinguer le comportement d'un régime de quench d'un régime de refroidissement. On remarque également que tous les préfacteurs sont interreliés en raison de leurs paramètres communs θ . Le modèle d'AP apprendra alors autant l'impact individuel que collectif des préfacteurs sur la prédiction $T_{\theta,i}^{n+1}$.

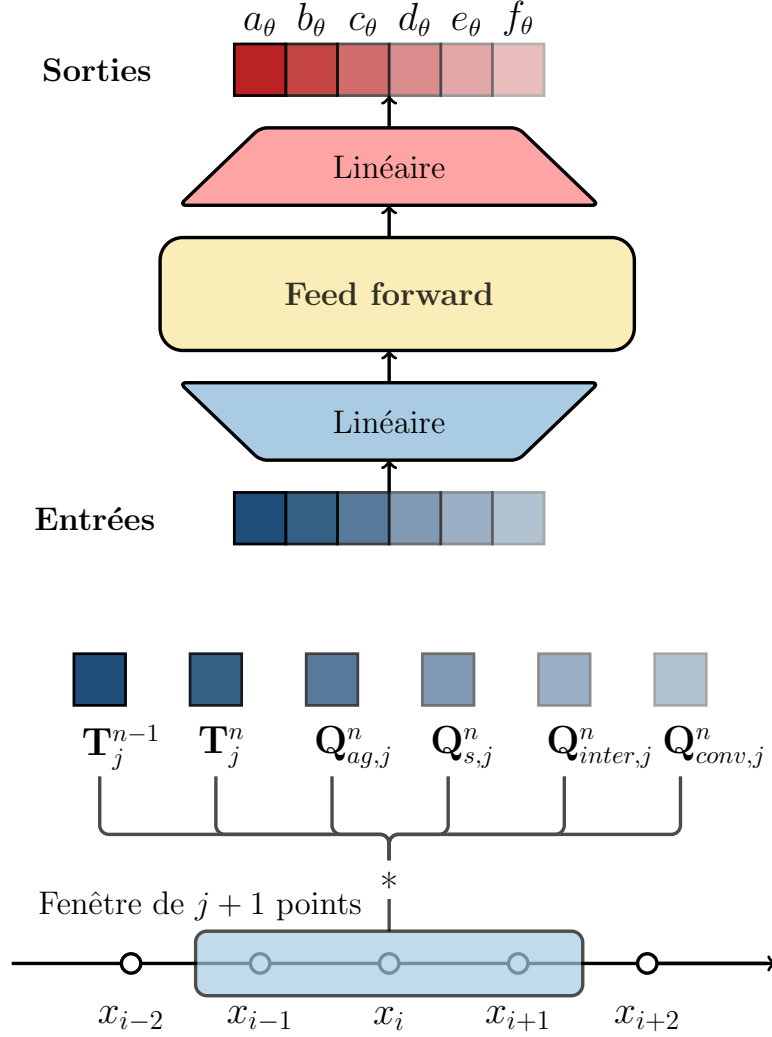


FIGURE 4.11 Schéma en blocs du modèle d'apprentissage profond *feed forward* utilisé dans le modèle RAPDi.

Ceci dit, le choix d'utiliser les variables thermiques en entrée n'est pas un choix anodin. Utiliser celles-ci au lieu des paramètres de simulation (R_{inter} , h_a , $I_{app}(t_n)$, σ_i , $\bar{A}_i$, $x_{c,i}$, n_D) permet au modèle d'apprentissage profond de raisonner en termes de chaleur sur $j + 1$ points, ainsi que de construire une interprétation abstraite et profonde du phénomène étudié. Cette représentation apprise a pour conséquence d'être plus naturelle au problème, résultant alors en des prédictions généralisables et moins propices au surapprentissage. En ce sens, la chaleur d'interface (Q_{inter}^n) permet de généraliser l'apprentissage pour différentes valeurs de R_{inter} . Les chaleurs de la couche d'argent et HTS permettent quant à elles de généraliser l'apprentissage pour différentes épaisseurs h_a , courants $I_{app}(t_n)$, régimes de stabilité et inhomogénéités de J_c . Finalement, la chaleur Q_{conv}^n aide de son côté à généraliser l'apprentissage pour les différents

régimes thermiques.

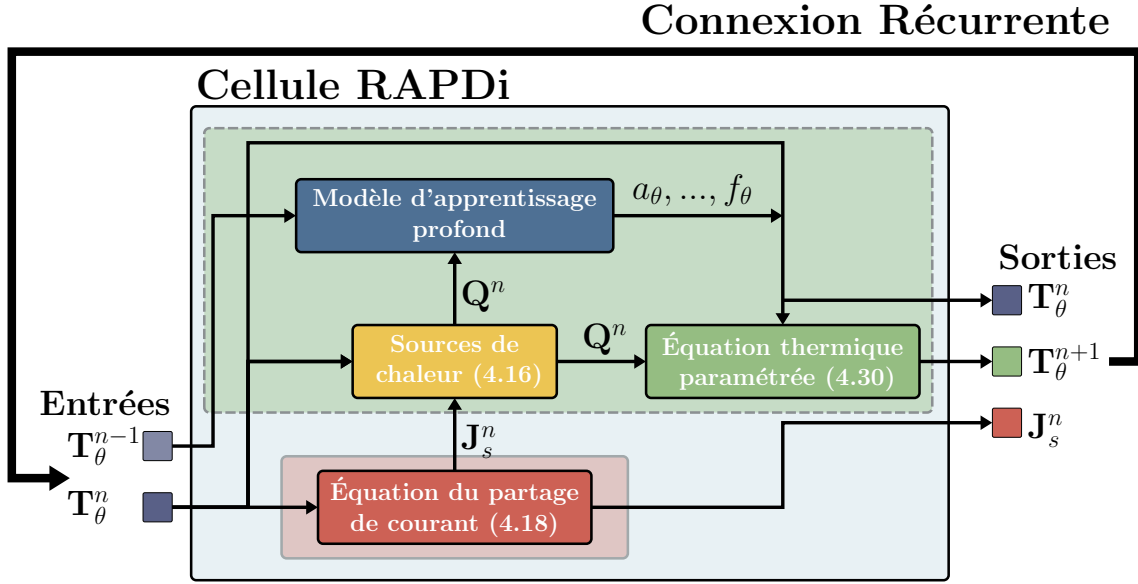


FIGURE 4.12 Schéma détaillé de la cellule récurrente RAPDi développé dans ce mémoire de maîtrise.

Le modèle RAPDi est donc formé du regroupement du problème électrique, du problème thermique ainsi que du modèle d'apprentissage profond présenté ci-dessus. Ces composantes constituent le coeur de la cellule récurrente du modèle RAPDi telle que présentée à la figure 4.12. La cellule RAPDi prend en entrée les températures T_θ^{n-1} et T_θ^n prédites précédemment et calcule le profil J_s^n à partir de l'équation de partage de courant (4.18). Ce résultat est alors combiné à T_θ^n afin de calculer chacune des sources de chaleur (équation (4.16)). Les sources de chaleur ainsi que les températures T_θ^{n-1} et T_θ^n sont alors combinées à l'entrée du modèle d'apprentissage profond afin de calculer $a_\theta, b_\theta, c_\theta, d_\theta, e_\theta$ et f_θ . Enfin, T_θ^{n-1} , T_θ^n , Q^n ainsi que tous les préfacteurs sont joints dans l'équation thermique paramétrée (4.30) afin d'obtenir la solution T_θ^{n+1} . Finalement, la connexion récurrente est appliquée : les sorties deviennent les entrées, c'est-à-dire que T_θ^{n+1} et T_θ^n deviennent respectivement T_θ^n et T_θ^{n-1} . Cette connexion récurrente est effectuée des centaines de fois de suite, i.e. pour chaque pas de temps de la simulation, afin d'obtenir des solutions complètes $J_s(x_i, t_n)$ et $T(x_i, t_n)$.

4.4.5 Propriété d'extension du domaine

Le modèle RAPDi est une méthode à maillage en raison de la présence sous-jacente de la méthode des différences finies. Toutefois, il n'est pas possible de modifier le maillage sans perte de généralité une fois l'entraînement terminé. Du moment que l'on considère une fenêtre de

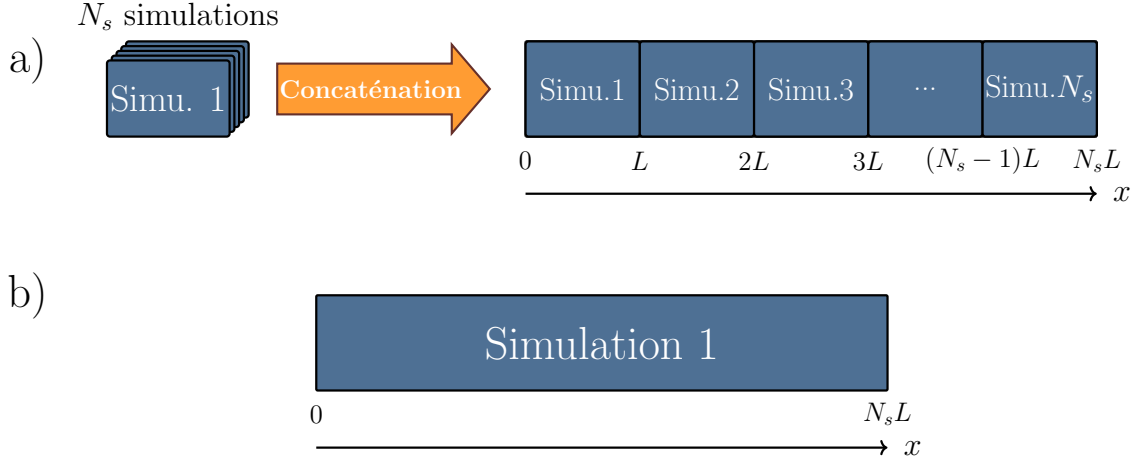


FIGURE 4.13 Construction d’une simulation d’un long ruban HTS (a) à partir d’une concaténation spatiale de N_s simulations et (b) la simulation résultante.

$j + 1$ points > 1 , le modèle d’apprentissage profond devient dépendant de la distribution locale des variables de chaleur et de températures ainsi que de l’espacement entre les nœuds du maillage Δx . Par ailleurs, les quantités Δx et Δt ont un impact direct sur la prédiction $\mathbf{T}_\theta^{n+1}$. La discrétisation spatiale et temporelle doivent donc conserver les mêmes valeurs que celles utilisées durant l’entraînement.

Cependant, le modèle RAPDi possède la propriété d’étendre son domaine contrairement à de nombreuses autres méthodes d’apprentissage profond. Cette propriété d’extension signifie qu’il est possible d’allonger le ruban HTS autant que souhaité post-entraînement, pour autant que l’on conserve la même discrétisation $(\Delta x, \Delta t)$, qu’on respecte les limites de l’espace des paramètres utilisées pour l’entraînement, et que des zones tampons sont présentes aux extrémités du ruban. La propriété d’extension du domaine émerge de l’invariance en translation lors de l’apprentissage : le modèle d’apprentissage profond apprend la physique seulement sur une fenêtre glissante de $j + 1$ points. En conséquence, allonger le ruban signifie seulement qu’on balaye cette fenêtre sur une plus longue distance, mais cela ne compromet en rien la dynamique apprise durant l’entraînement. Cette propriété est également vérifiée sur la dimension temporelle : le modèle n’a jamais vu la longueur totale L ou le temps final $t_f = 10$ ms, seulement la discrétisation $(\Delta x, \Delta t)$. Notons que cette notion de fenêtre glissante est en réalité équivalente à une couche de convolution 1-D d’un réseau de neurones convolutif (CNN) où un seul filtre de taille $j + 1$ éléments est utilisé.

Dans le but de vérifier les performances du modèle RAPDi sur de longs rubans, une simulation d’un long ruban sera fabriquée à partir d’une concaténation spatiale de N_s simulations

de ruban de longueur L , tel qu'illustré à la figure 4.13. Cette méthodologie permet d'obtenir des rubans de plusieurs mètres de long, ce qui serait impraticable autrement, et constitue d'ailleurs une limite des modèles éléments finis actuels. En ce sens, un modèle entraîné sur des rubans d'une longueur L pourra prédire la dynamique d'un ruban d'une longueur $N_s \times L$ et comparer la solution obtenue avec celle de la MEF. Il convient de noter que la solution concaténée issue de la MEF ne peut être utilisée comme simulation de référence que si les N_s simulations sont indépendantes. Dans le cas contraire où un quench survient à $x = L$, alors la solution par éléments finis concaténée ne serait plus exacte. En ce sens, l'approche RAPDi permet d'augmenter significativement la portée des simulations numériques actuellement accessibles.

4.4.6 Stratégies d'entraînement

Les solutions discrétisées $T(x_i, t_n)$ et $J_s(x_i, t_n)$ issues des simulations par éléments finis sont représentées par des matrices de taille $N_x \times N_t$ où N_t est le nombre de points temporels. Toutefois, comme N_t est dans l'ordre de la centaine, il est compliqué d'entraîner le modèle proposé sur une séquence de N_t états par souci de temps et de mémoire. En ce sens, chaque simulation est plutôt subdivisée en n_s séquences temporelles de longueur $l_s \geq 3$ afin d'entraîner le modèle RAPDi. Les deux premiers états de la séquence représentent les instants t_{n-1} et t_n respectivement.

La fonction de perte $\mathcal{L}$ peut alors considérer la longueur l_s lors de l'entraînement par le biais d'une moyenne. Dans le cadre de ce mémoire, la fonction de perte utilisée est

$$\mathcal{L}(\theta) = \frac{1}{n_s} \sum_{s=1}^{n_s} \ell_s(\mathbf{T}_\theta^n, \mathbf{T}^n), \quad (4.34)$$

avec

$$\ell_s(\mathbf{T}_\theta^n, \mathbf{T}^n) = \frac{1}{l_s - 2} \sum_{l=1}^{l_s-2} \left\| \mathbf{T}_\theta^{n+l} - \mathbf{T}^{n+l} \right\|_2, \quad (4.35)$$

où l'équation (4.34) représente l'erreur euclidienne moyenne sur l'ensemble de toutes les séquences entre les prédictions du modèle RAPDi $\mathbf{T}_\theta$ et les solutions de la MEF $\mathbf{T}$. Notons cependant qu'il est peu fréquent de sommer sur n_s lorsqu'on possède beaucoup de données pour des raisons de mémoire. En pratique, on définit plutôt un mini-lot de taille $\mathcal{B} \ll n_s$, et on calcule alors l'équation (4.34) suivi d'une rétropropagation à plusieurs reprises pour une époque donnée. Par ailleurs, une stratégie adaptative est implémentée afin d'entraîner le modèle RAPDi sur de longues séquences : le modèle est pré-entraîné pour $l_s = 3$ jusqu'à une tolérance désirée, et l_s est ensuite augmenté progressivement à chaque fois que la tolérance fixée est ré-atteinte. Cette stratégie permet donc au modèle d'apprendre plus efficacement en

commençant par apprendre de petites séquences et en augmentant la difficulté graduellement. Ceci dit, l'entraînement du modèle RAPDi reste encore trop coûteux en calcul en raison de la nature implicite de l'équation du partage de courant. Il est impraticable à l'échelle de ce mémoire de résoudre l'équation (4.18) pour toutes les séquences du jeu de données, et ce pour toutes les époques (nombre d'itérations où le modèle d'apprentissage profond a vu l'entièreté de l'ensemble d'entraînement). En ce sens, une stratégie de forçage par enseignant est donc adoptée afin de pré-calculer toutes les sources thermiques de l'équation (4.16). De cette manière, les calculs associés au problème électrique sont évités, et le problème thermique peut être calculé à faible coût, grâce à la formulation explicite de l'équation (4.30). L'équation du partage de courant est bien évidemment calculée une fois l'entraînement terminé. Le processus d'optimisation des paramètres θ (équation (4.34)) dépend donc seulement de la température, négligeant le problème électrique lors de l'entraînement. Notons qu'il est possible d'élaborer une stratégie alternative où le modèle est pré-entraîné à partir du forçage par enseignant, et ensuite entraîné en résolvant l'équation (4.18). Bien que cette dernière soit plus précise, ses coûts élevés en calcul restent un désavantage important.

Enfin, il est important de mentionner que les valeurs numériques d'entrées du modèle d'apprentissage profond peuvent varier de 0 jusqu'à 10^7 . Il est alors essentiel de normaliser celles-ci afin d'éviter tout problème de stabilité numérique ou tout phénomène de pré-sélection des variables. En ce sens, chacune des entrées du modèle d'apprentissage profond subit une normalisation min-max

$$Q \leftarrow \frac{Q - \min(Q)}{\max(Q) - \min(Q)} \quad (4.36)$$

durant l'entraînement. Il en va de même pour les variables associées à la température.

CHAPITRE 5 MODÉLISATION DU QUENCH DANS DES RUBANS D'ARCHITECTURE 2-D UNIFORME

Maintenant que le modèle RAPDi a été introduit, ce chapitre démontre la validité de l'approche en modélisant des rubans HTS uniformes 2-D à partir du modèle proposé. On entame ce chapitre avec l'optimisation de l'architecture du modèle et la caractérisation de ses performances sur certaines métriques. Une fois le meilleur modèle obtenu, des exemples de solutions thermiques et électriques sont présentés. Le chapitre se conclut avec la prédiction du comportement en quench d'un ruban long de plus d'un mètre.

5.1 Optimisation de l'architecture et des performances du modèle RAPDi

L'architecture exacte du modèle d'apprentissage profond sous-jacent au modèle RAPDi est impossible à déterminer avant l'entraînement. Celle-ci dépend de la complexité de la tâche d'apprentissage, et donc du jeu de données. Il est ainsi nécessaire d'étudier un espace varié d'architectures et d'hyperparamètres afin d'identifier le modèle d'apprentissage profond le plus performant pour le jeu de données à notre disposition.

Dans cette optique, le jeu de données utilisé dans ce chapitre est constitué de 103 simulations électrothermiques 2-D résolues par la méthode des éléments finis pour un ruban uniforme, tel que présenté à la section 4.2.2. La MEF résout le problème avec un maillage $\Delta x = 70 \text{ } \mu\text{m}$ ($N_x = 1001$) et un pas de temps $\Delta t = 10 \text{ } \mu\text{s}$ ($N_t = 1001$). Les simulations sont réparties en ensembles d'entraînement, de validation et de test, avec 80, 2 et 21 simulations respectivement. Chaque ensemble possède autant de simulations en régime dominé par les pertes Joule que de simulations dominées par le refroidissement.

De son côté, le modèle RAPDi est implémenté en PyTorch. Plusieurs de ses hyperparamètres sont maintenus fixes, dont les fonctions d'activation (ReLU), la taille de mini-lot $\mathcal{B}$ (32), le nombre d'époques (200) et le taux d'apprentissage ($7.5 \cdot 10^{-5}$), afin d'alléger l'espace de modèles à étudier. Les paramètres θ du modèle sont optimisés à partir de la méthode d'ADAM (section 3.1.4). Le modèle RAPDi est entraîné avec un maillage $\Delta x = 140 \text{ } \mu\text{m}$ ($N_x = 501$) et un pas de temps $\Delta t = 20 \text{ } \mu\text{s}$ ($N_t = 501$), soit une discrétisation deux fois plus grossière que la MEF.

TABLEAU 5.1 Modèles RAPDi étudiés à base de réseau *feed forward*.

Nom du modèle	Architecture	Dropout p	Taille fenêtre glissante
NN-Mini	[6,25,6]	0.05	1
NN-Petit-1	[6,75,6]	0.05	1
NN-Petit-2	[6,75,75,6]	0.05	1
NN-Petit-3	[18,75,75,6]	0.05	3
NN-Moyen-1	[18,150,150,6]	0.05	3
NN-Moyen-2	[30,200,200,200,6]	0.05	5
NN-Moyen-3	[30,250,250,250,250,6]	0.075	5
NN-Gros-1	[30,250,250,250,250,250,6]	0.1	5
NN-Gros-2	[42,300,300,300,300,6]	0.1	7
NN-Gros-3	[54,400,400,400,400,6]	0.2	9

TABLEAU 5.2 Modèles RAPDi étudiés à base d'apprentissage profond avec attention.

Nom du modèle	N_B	Architecture FF	Dropout p	Taille fenêtre glissante
APA-Profond-1	6	[30,30]	0.05	5
APA-Profond-2	4	[30,30,30]	0.075	5
APA-Petit-1	1	[100,100]	0.05	3
APA-Petit-2	1	[100,100,100]	0.075	3
APA-Moyen-1	1	[200,200]	0.05	5
APA-Moyen-2	2	[200,200]	0.05	5
APA-Moyen-3	2	[200,200,200]	0.05	5
APA-Gros-1	2	[250,250,250]	0.075	5
APA-Gros-2	2	[250,250,250]	0.1	7
APA-Gros-3	3	[250,250,250]	0.15	5

5.1.1 L'espace de modèles RAPDi étudié

Les hyperparamètres variés entre chaque modèle RAPDi étudié sont l'architecture du modèle d'AP (nombre de neurones, nombre de couches profondes, présence de mécanismes d'attention), la probabilité de *dropout* p , ainsi que la taille de la fenêtre glissante. Dans le cadre de modèles RAPDi à base d'un réseau *feed forward*/réseaux de neurones (figure 4.11), on étudie uniquement 11 architectures afin de restreindre l'espace de modèles étudié. Ces diverses architectures sont présentées au tableau 5.1. La notation [18,75,75,6] désigne le nombre de neurones à chacune des couches. Notons par ailleurs que la valeur p au tableau 5.1 est d'autant plus grande que le modèle est susceptible au surapprentissage.

On introduit également un second type de modèle d'apprentissage profond incorporant des mécanismes d'attention à une seule tête et une connexion résiduelle, tel qu'illustré à la fi-

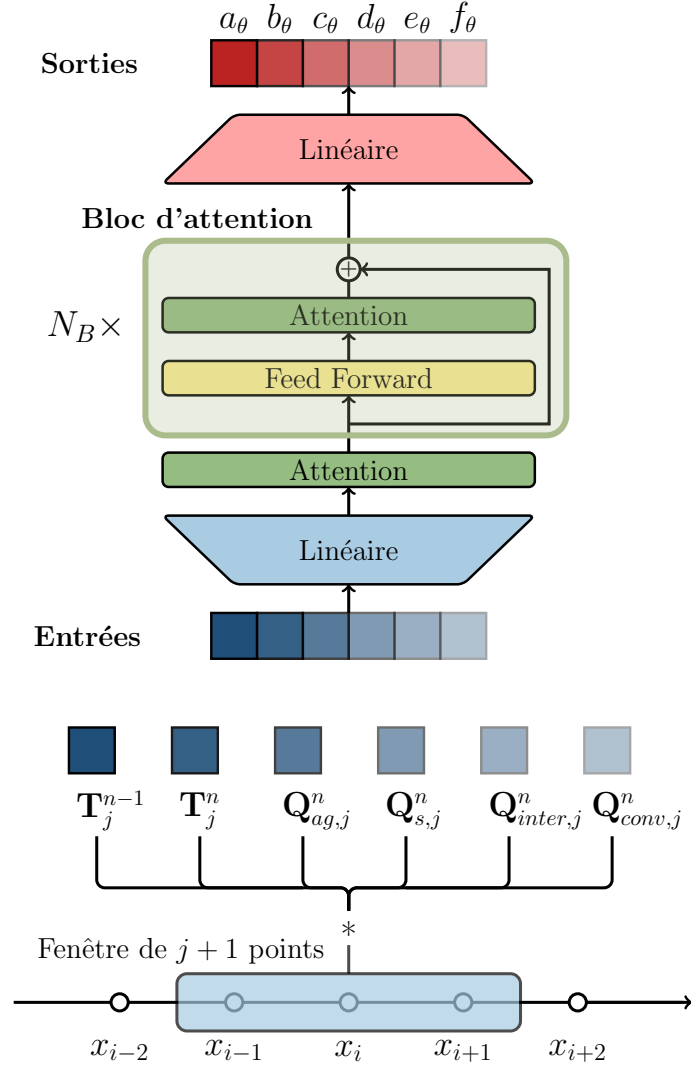


FIGURE 5.1 Modèle d'apprentissage profond avec attention (APA). Le modèle est formé de N_B blocs d'attention.

gure 5.1. On réfère cette architecture à un modèle d'apprentissage profond avec attention (APA). On étudie pour cette dernière 10 architectures différentes, telles que présentées au tableau 5.2. Notons que la colonne « Architecture FF » fait référence à l'architecture du réseau *feed forward* de la figure 5.1, et la colonne « N_B » fait référence au nombre de blocs d'attention utilisés.

5.1.2 Performances selon les métriques

Les performances de chaque modèle listé aux tableaux 5.1 et 5.2 sont quantifiées à l'aide de la norme euclidienne de l'erreur $e_T(t)$ associée à la variable T telle que

$$e_T(t_n) = \frac{1}{N_s} \sum_{s=1}^{N_s} \left\| \mathbf{T}_\theta^n - \mathbf{T}^n \right\|_2. \quad (5.1)$$

L'équation (5.1) calcule donc l'erreur associée à $\mathbf{T}_\theta^n$ sous la norme L_2 dans l'espace à chaque instant t_n et la moyenne est prise sur l'ensemble des simulations. L'erreur euclidienne moyenne en température $\langle e_T \rangle$, en densité de courant $\langle e_{J_s} \rangle$ ainsi qu'en potentiel électrique $\langle e_{V_s} \rangle$ serviront alors de critère de sélection afin de déterminer le modèle le plus performant. Ces trois métriques permettront d'avoir un aperçu autant sur la solution thermique que sur les solutions électriques. Notons par ailleurs que, comme le potentiel électrique V_s s'exprime tel que

$$V_s(x, t) = \int_x^L \rho_s(x', t) J_s(x', t) dx' + V_0, \quad (5.2)$$

où V_0 est une constante d'intégration, V_s est particulièrement sensible à l'accumulation d'erreurs en raison de l'intégration de la résistivité électrique ρ_s et de la densité de courant J_s , deux quantités elles-mêmes inexactes. L'erreur associée à $V_s(x, t)$ est donc une métrique informative afin de déterminer la fiabilité et l'exactitude d'un modèle.

Les erreurs mesurées pour les modèles RAPDi à base de réseau *feed forward* et d'APA sont présentées aux tableaux 5.3 et 5.4 respectivement. Les performances du modèle Levin *et al.* (modèle de référence) sont également présentées à des fins comparatives. Notons toutefois que les résultats obtenus ne sont pas déterministes en raison de l'initialisation aléatoire des paramètres θ ainsi que de l'ordre des lots lors de l'entraînement.

Soulignons que la majorité des modèles proposés produisent une erreur $\langle e_T \rangle$, $\langle e_{J_s} \rangle$ et $\langle e_{V_s} \rangle$ 10 fois inférieures à celle du modèle de référence, tel que présenté dans les tableaux 5.3 et 5.4. Du côté des modèles RAPDi à base de réseau *feed forward*, le modèle ayant les meilleures performances est le modèle NN-Moyen-3, avec des erreurs $\langle e_T \rangle$, $\langle e_{J_s} \rangle$ et $\langle e_{V_s} \rangle$ de 5.83 K, 0.192 MA cm⁻² et 0.137 V respectivement. Celui-ci surpasse tous les autres modèles pour chacun des critères sélectionnés. Notons cependant que le modèle NN-Gros-1 reste un candidat avec des performances comparables. On remarque également que les deux modèles les plus performants de cette catégorie utilisent chacun une fenêtre glissante d'une taille de 5 éléments.

De l'autre côté, le modèle RAPDi à base d'APA le plus performant est le modèle APA-Moyen-2 avec des erreurs $\langle e_T \rangle$, $\langle e_{J_s} \rangle$ et $\langle e_{V_s} \rangle$ de 6.79 K, 0.210 MA cm⁻² et 0.107 V respectivement.

TABLEAU 5.3 Erreurs des modèles RAPDi étudiés à base de réseau *feed forward*.

Nom du modèle	$\langle e_T \rangle$ (K)	$\langle e_{J_s} \rangle$ (MA cm ⁻²)	$\langle e_{V_s} \rangle$ (V)
NN-Mini	8.25	0.306	0.163
NN-Petit-1	8.17	0.298	0.167
NN-Petit-2	8.29	0.295	0.196
NN-Petit-3	6.68	0.230	0.165
NN-Moyen-1	6.98	0.245	0.178
NN-Moyen-2	7.80	0.293	0.204
NN-Moyen-3	5.83	0.192	0.137
NN-Gros-1	6.1	0.209	0.147
NN-Gros-2	7.68	0.301	0.232
NN-Gros-3	7.23	0.276	0.197
Modèle de référence	74.1	2.19	2.41

TABLEAU 5.4 Erreurs des modèles RAPDi étudiés à base d'APA.

Nom du modèle	$\langle e_T \rangle$ (K)	$\langle e_{J_s} \rangle$ (MA cm ⁻²)	$\langle e_{V_s} \rangle$ (V)
APA-Profond-1	6.25	0.229	0.149
APA-Profond-2	11.7	0.492	0.374
APA-Petit-1	7.82	0.268	0.204
APA-Petit-2	7.60	0.281	0.217
APA-Moyen-1	7.95	0.243	0.202
APA-Moyen-2	6.79	0.210	0.107
APA-Moyen-3	6.62	0.220	0.140
APA-Gros-1	8.25	0.317	0.250
APA-Gros-2	7.66	0.260	0.190
APA-Gros-3	7.10	0.252	0.126
Modèle de référence	74.1	2.19	2.41

Bien que ce ne soit pas le modèle avec l'erreur $\langle e_T \rangle$ la moins élevée, ses faibles erreurs associées aux solutions électriques le distinguent des autres architectures proposées.

On remarque que les performances des modèles RAPDi à base de réseau *feed forward* et d'APA sont plutôt similaires. Intégrer un mécanisme d'attention revient finalement à un réglage fin de l'architecture. Le mécanisme d'attention ajoute une complexité parfois superflue au modèle, ce qui est reflété sur les métriques d'évaluation : les performances en température et densité de courant du modèle NN-Moyen-3 surpassent celles du modèle APA-Moyen-2.

Par ailleurs, la figure 5.2 illustre les performances de ces deux modèles ainsi que les performances d'un modèle de coefficients effectifs (MCE) durant l'entraînement (figure 5.2a) et sur l'ensemble de test (figure 5.2b). On constate à la figure 5.2a que les fonctions de perte des modèles RAPDi à base de réseau *feed forward* et d'APA semblent se superposer, souvent

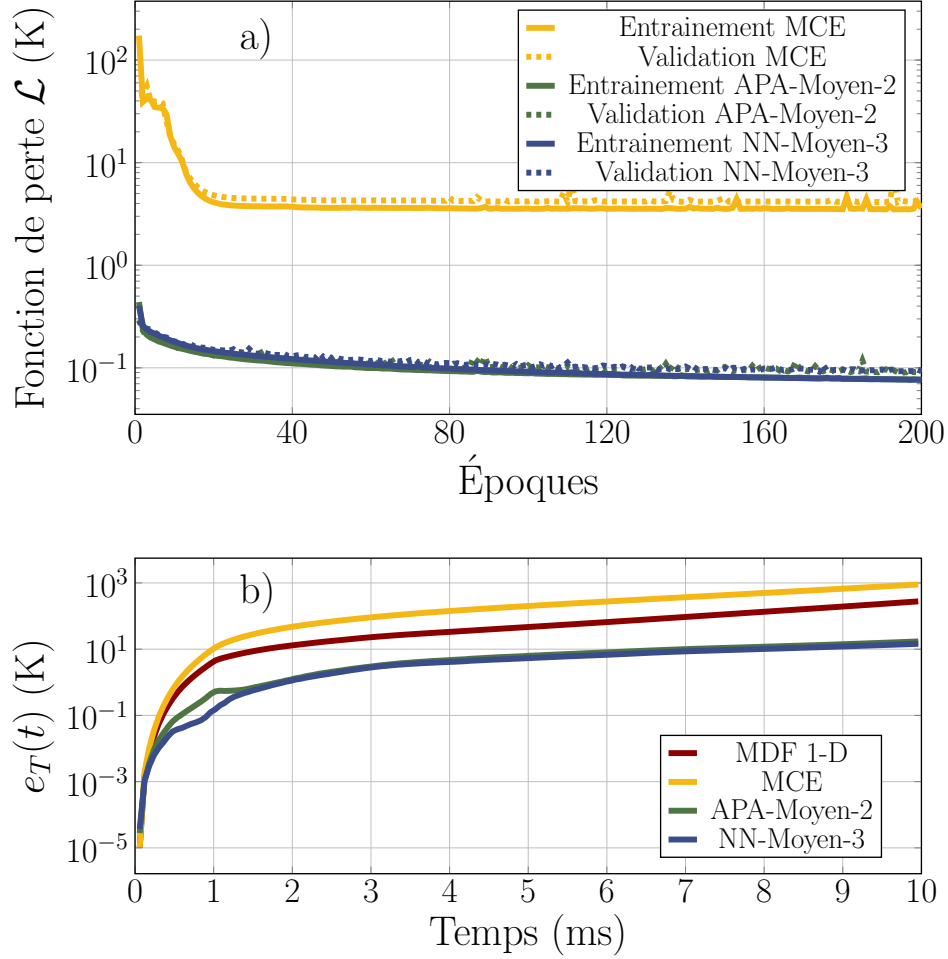


FIGURE 5.2 (a) Courbes d'apprentissage des modèles RAPDi NN-Moyen-3, APA-Moyen-2 et MCE durant l'entraînement ; (b) prédictions avec ces mêmes modèles sur l'ensemble de test pour des rubans uniformes.

synonyme de performances similaires. De l'autre côté, la fonction de perte du MCE est bien plus élevée, suggérant des performances médiocres comparativement aux deux autres modèles. C'est notamment ce qu'on observe à la figure 5.2b, où l'erreur du MCE est supérieure à celle du modèle de référence (MDF 1-D). Tel que discuté à la section 4.4.1, le MCE est incapable de représenter deux régimes de stabilité distincts simultanément. Cela est d'autant plus vrai lorsque les paramètres géométriques (R_{inter} et h_a) sont variés. Les erreurs des modèles RAPDi à base de *feed forward* et d'APA sont quant à elles très similaires, hormis autour de 1 ms. Le modèle NN-Moyen-3 semble mieux reproduire la dynamique lors de la nucléation du quench, ce qui explique pourquoi les métriques favorisent ce modèle. La clé d'un apprentissage réussi réside souvent dans la précision du modèle lors de la nucléation du quench, minimisant alors l'accumulation subséquente d'erreur.

TABLEAU 5.5 Temps de calcul moyen d’une simulation et le facteur d’accélération obtenu par rapport à la MEF 2-D pour différents modèles.

Modèle	Temps de calcul	Facteur d’accélération ($\times$)
MEF 2-D	1.5 h	1
MDF 1-D	1.5 s	3600
RAPDi 1-D	12 s	450

5.1.3 Performances temporelles

Les temps de calcul d’une simulation effectuée par la méthode des éléments finis 2-D, par le modèle de référence et par le modèle RAPDi (NN-Moyen-3) est présenté au tableau 5.5.

Pour des rubans uniformes de taille $L = 7$ cm, le modèle RAPDi est en moyenne 450 fois plus rapide que la MEF 2-D, alors que le facteur d’accélération s’élève à 3600 pour la MDF 1-D. Le modèle RAPDi est toutefois 10 fois plus performant sur les métriques choisies que le modèle de référence. Bien que l’accélération des calculs s’explique majoritairement par la formulation unidimensionnelle, une partie de celle-ci s’explique également par la discrétisation deux fois plus grossière du modèle RAPDi et de la MDF 1-D par rapport à la MEF 2-D. Par ailleurs, la différence majeure entre le temps de calcul du modèle RAPDi (12s) et la MDF 1-D (1.5s) est expliquée par la présence du modèle d’apprentissage profond. Certaines modifications pourraient toutefois être mises en oeuvre au niveau des librairies utilisées afin d’augmenter le facteur d’accélération du modèle RAPDi au-delà de 1000.

La durée de l’entraînement de la partie thermique du modèle RAPDi sur GPU (GeForce RTX 3060 Ti) varie entre 60 et 90 minutes selon le nombre de paramètres du modèle d’apprentissage profond utilisé. Le même entraînement peut prendre jusqu’à dix fois ce temps si le modèle est exécuté sur un CPU conventionnel.

5.2 Solutions en température d’un ruban uniforme

Les solutions thermiques $T(x, t)$ de la MEF 2-D, de la MDF 1-D (modèle de référence) et du modèle RAPDi sont présentées dans le cadre de deux régimes de stabilité distincts. Rappelons que, dans le cas de la MEF 2-D, la solution est échantillonnée au centre de la couche supraconductrice. Le modèle de RAPDi prédit cette même solution. Le modèle de référence obtient quant à lui la température d’un ruban uniforme aux propriétés thermiques homogènes. Notons que toutes les simulations présentées dans les prochaines sous-sections sont issues de l’ensemble de test, et le modèle RAPDi retenu est le modèle NN-Moyen-3.

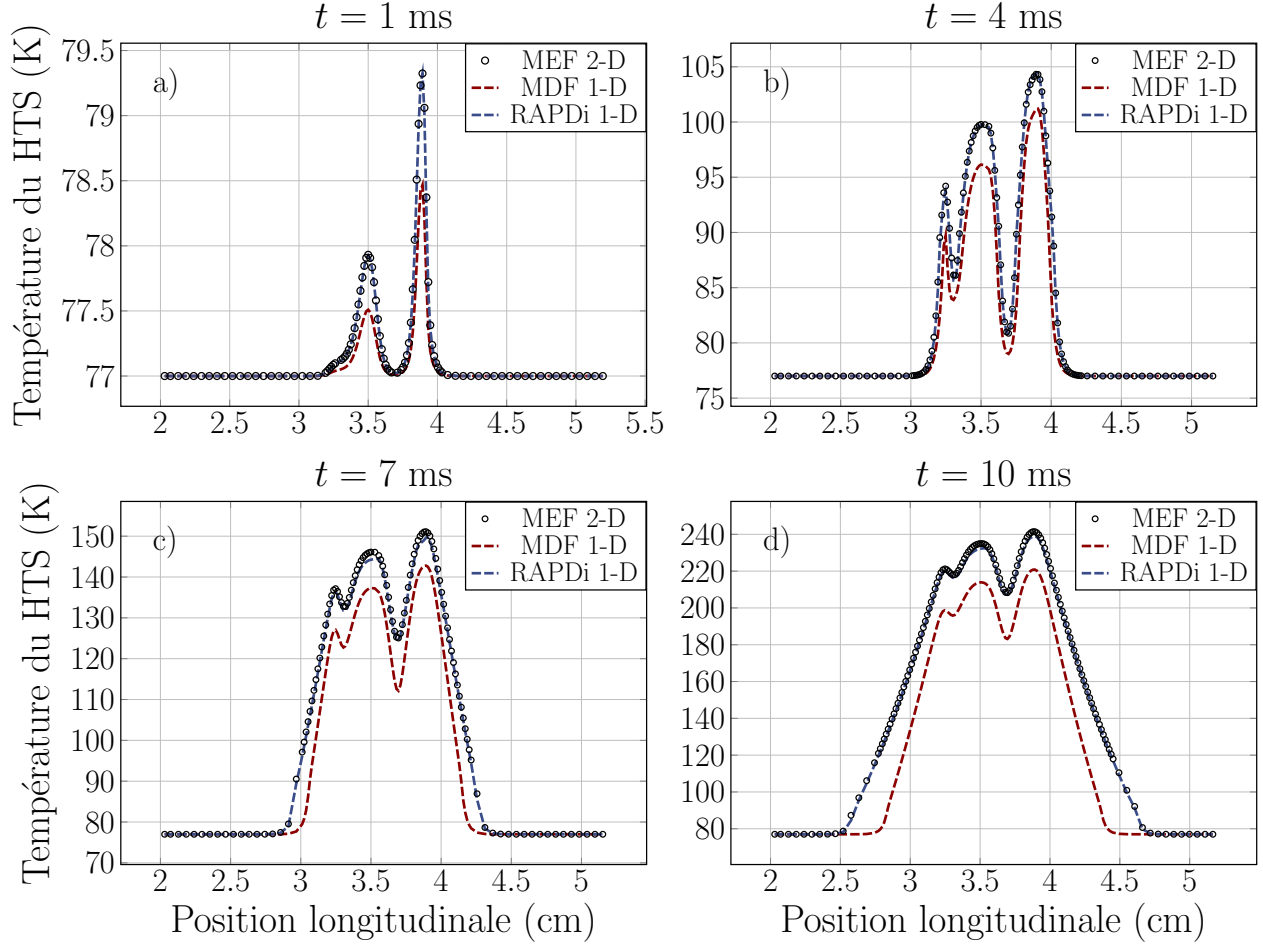


FIGURE 5.3 Température longitudinale à l'intérieur de la couche HTS d'un ruban uniforme aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms lors d'un régime dominé par les pertes par effet Joule pour la MEF 2-D, la MDF 1-D (modèle de référence) et le modèle RAPDi. La résistance d'interface et l'épaisseur de la couche d'argent sont de $6.7 \mu\Omega \text{ cm}^2$ et $1.7 \mu\text{m}$, respectivement.

5.2.1 Solution thermique pour le régime dominé par les pertes Joule

La figure 5.3 illustre une solution thermique de l'ensemble de test pour un régime dominé par les pertes par effet Joule à différents temps, où la résistance d'interface $R_{inter} = 6.7 \mu\Omega \text{ cm}^2$ et l'épaisseur d'argent $h_a = 1.7 \mu\text{m}$. On remarque que la solution obtenue par le modèle RAPDi s'aligne étroitement avec la solution de la MEF 2-D, et ce, pour des résistances d'interface et des épaisseurs d'argent encore jamais vues par l'entraînement. L'erreur devrait également être maximale au temps final d'un point de vue des méthodes numériques classiques, en raison de la température strictement croissante et de la nature récurrente du problème. Malgré cela, la solution du modèle RAPDi au temps final t_f (figure 5.3d) reste exacte. Ceci dit, gardons

en tête qu'il est toutefois possible d'observer une erreur $e_T(t)$ qui n'est pas nécessairement strictement croissante. Des fluctuations de $e_T(t)$ peuvent survenir du fait que le modèle RAPDi pondère la chaleur durant la simulation. C'est notamment ce qu'on observe à la figure 5.2 vers 1 ms lorsqu'on transite du régime supraconducteur au régime normal (autour de T_c), une transition parfois difficile à apprendre pour le modèle.

Du côté du modèle de référence, on remarque une disparité grandissante au fil du temps entre sa solution et celle de la MEF 2-D. Cette disparité s'explique par le fait que la MDF 1-D sous-estime la température lors de la nucléation du quench, tel qu'illustré aux figures 5.3a et 5.3b. Cette sous-évaluation de la température a pour conséquence d'initier la zone normale en retard, ce qui entraîne alors une propagation tardive de celle-ci, et donc, une zone normale déphasée par rapport à la MEF 2-D, tel qu'illustré aux figures 5.3c et 5.3d. Le modèle de référence semble toutefois bien approximer la NZPV, mais le retard accumulé rend sa solution complètement inexacte. Le modèle de Levin *et al.* est donc mal adapté aux applications qui requièrent une modélisation précise de la nucléation de quench, telles que les applications d'aimants supraconducteurs ou de limiteurs de courant de court-circuits.

Notons par ailleurs que le modèle de référence sous-estime la température en raison de l'hypothèse d'homogénéisation sous-jacente. Lors de la nucléation du quench, les propriétés thermiques au coeur de la physique sont largement dominées par celles de la couche supraconductrice et stabilisatrice en raison du courant de transport qui y circule ainsi que de leur faible épaisseur (voir tableau 4.1). Pour une échelle de temps de 10 ms, la chaleur n'a pas le temps de se dissiper à l'intérieur du substrat. En ce sens, si l'on utilise malgré tout une hypothèse d'homogénéisation, on ajoute artificiellement de la masse thermique dans ces couches minces, ce qui fait dévier la solution de la réalité. La génération de chaleur, et donc la température obtenue, est évidemment plus faible que la MEF 2-D pour une même densité de courant. Ce résultat était attendu puisque le critère de chaleur surfacique de 10^4 W/m² mentionné à la section 2.4.2 est longuement dépassé. Le modèle de Levin *et al.* atteint une chaleur surfacique jusqu'à 10^7 W/m².

5.2.2 Solution thermique pour le régime dominé par le refroidissement

La figure 5.4 illustre un exemple d'une simulation lors d'un régime dominé par le refroidissement, où $R_{inter} = 7.6 \mu\Omega \text{ cm}^2$ et $h_a = 1.9 \mu\text{m}$. On observe que le modèle RAPDi concorde très bien avec la méthode des éléments finis 2-D pour tous les instants. Toutefois, on remarque que la solution à l'instant $t = 4$ ms est légèrement surestimée tandis que celle-ci redevient exacte à $t = 10$ ms. Ce phénomène est simplement une conséquence directe de pondérer la chaleur comme expliqué précédemment. Par ailleurs, notons qu'entraîner le modèle RAPDi sur les

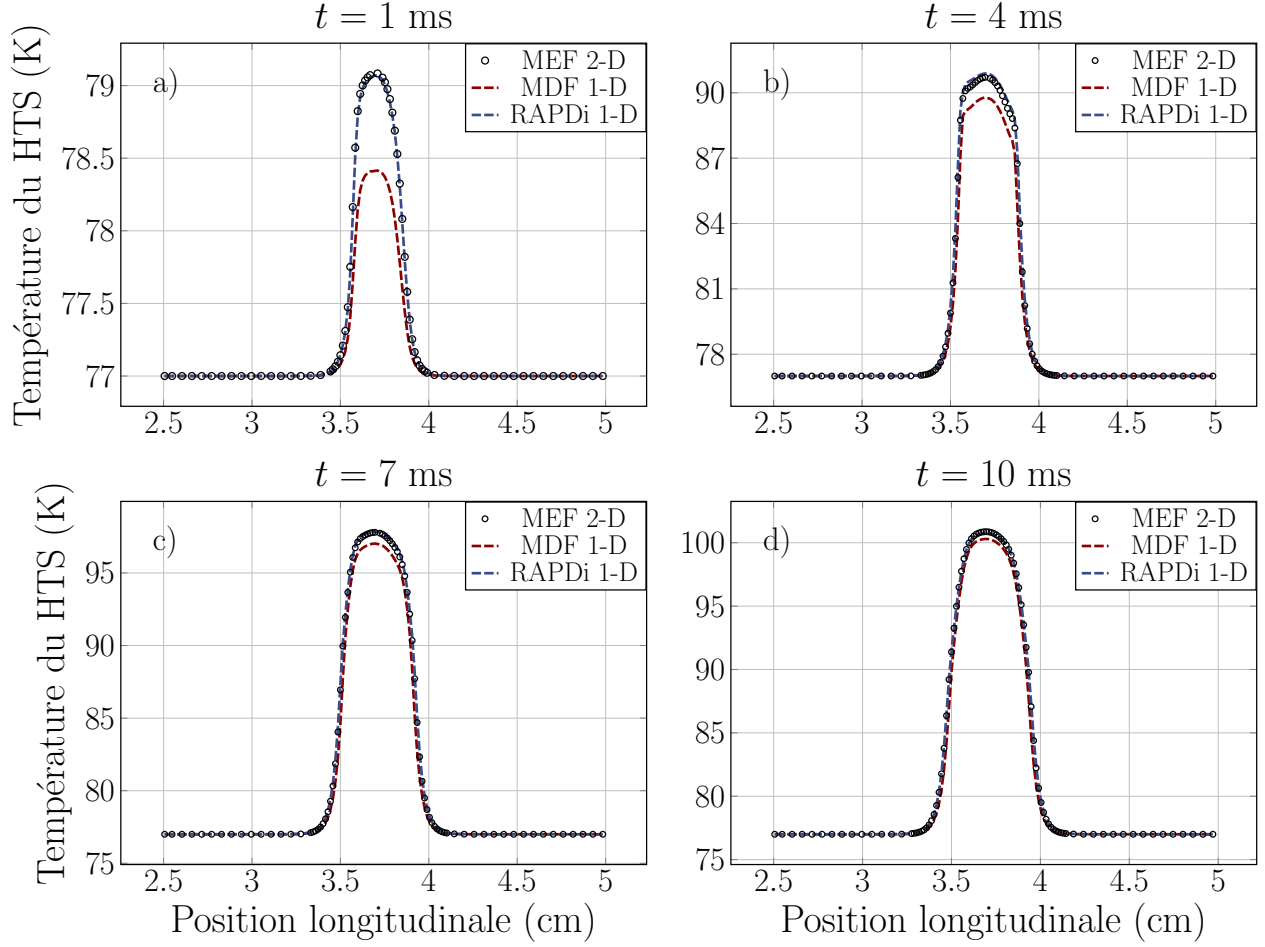


FIGURE 5.4 Température longitudinale à l'intérieur de la couche HTS aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms lors d'un régime dominé par le refroidissement pour la MEF 2-D, la MDF 1-D (modèle de référence) et le modèle RAPDi. La résistance d'interface et l'épaisseur de la couche d'argent sont de $7.6 \mu\Omega \text{ cm}^2$ et $1.9 \mu\text{m}$, respectivement.

deux régimes de stabilité simultanément peut rendre l'apprentissage des solutions en régime de refroidissement plus ardu. Plus précisément, la température moyenne est généralement plus faible lors d'un régime dominé par le refroidissement et donc, la pénalisation associée (équation (4.34)) l'est également. Le modèle concentre ainsi davantage son apprentissage sur les régimes dominés par les pertes Joule où la pénalisation est plus sévère. On peut toutefois mitiger ces effets en utilisant une fonction de perte relative, ou en pondérant la fonction de perte selon la température ($\lambda_R(T)\mathcal{L}$).

Du côté du modèle de référence, on observe une bien meilleure performance que dans le cas d'un régime dominé par les pertes par effet Joule. Tout d'abord, on remarque une sous-estimation de la température à la figure 5.4a, ce qui s'explique par l'initialisation du quench

aux premiers instants où une importante chaleur est générée. Cependant, la solution de la MDF 1-D semble « s'améliorer » pour $t > 1$ ms. Cela est expliqué par le fait que le courant est drastiquement diminué à partir de 1 ms (voir figure 4.8) pour un régime dominé par le refroidissement tel que défini à la section 4.2.2. L'hypothèse d'homogénéisation devient alors de plus en plus valide au fil du temps. De plus, comme la zone normale ne se propage pas, la solution déphasée « rattrape » la MEF 2-D. Celle-ci reste toutefois toujours inférieure à la véritable solution, contrairement au modèle RAPDi proposé. Ce dernier permet donc d'obtenir $T(x, t)$ au centre de la couche supraconductrice lors d'un emballement thermique, là où la MDF 1-D échoue, en plus d'être plus précis que celle-ci dans le cas d'un régime dominé par le refroidissement. Celui-ci reproduit également la solution thermique de la MEF 2-D avec une discrétisation deux fois plus grossière.

Ceci dit, les simulations présentées restent seulement deux exemples parmi 21 simulations distinctes. Pour une vue d'ensemble, la figure 5.5 illustre le comportement de $e_T(t)$ pour les deux régimes de stabilité thermique. On remarque que malgré les fluctuations, l'erreur associée au régime dominé par les pertes Joule présente une tendance croissante, tandis que l'erreur correspondant au régime dominé par le refroidissement stagne après un certain temps. L'erreur est stable dans ce deuxième cas puisque la température ne varie pratiquement plus.

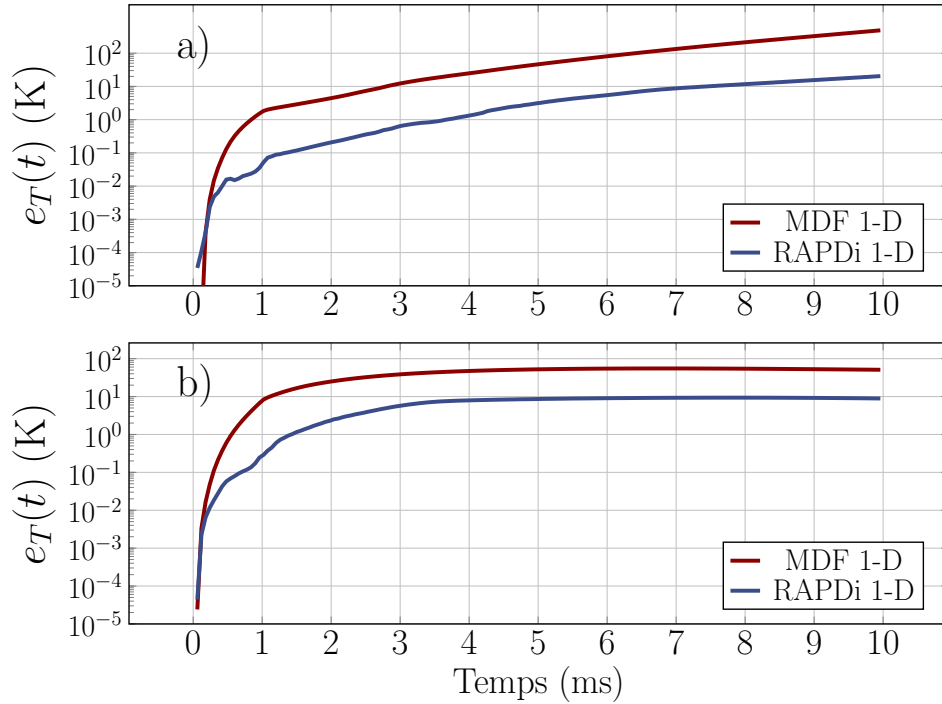


FIGURE 5.5 Erreur euclidienne moyenne $e(t)$ associée (a) aux régimes dominés par les pertes Joule et (b) aux régimes dominés par le refroidissement.

Pour le régime dominé par les pertes Joule, le modèle RAPDi possède une erreur 24 fois plus faible à $t = 10$ ms comparativement au modèle de référence. Ce facteur n'est que 6 dans le cas de simulations dominées par le refroidissement. Rappelons également que le modèle n'est entraîné que sur une séquence temporelle l_s de quelques états, tandis que les résultats obtenus aux figures 5.3 et 5.4 sont issus de prédictions sur une séquence temporelle de 501 états ($N_t = 501$).

5.3 Solutions électriques d'un ruban uniforme

Les solutions électriques $J_s(x, t)$ et $V_s(x, t)$ associées à la solution thermique de la figure 5.3 seront présentées dans les prochaines sous-sections. Les solutions électriques correspondant au régime dominé par le refroidissement (figure 5.4) ne sont quant à elles pas présentées puisque la solution $T(x, t)$ varie très peu, et par conséquent, J_s et V_s aussi.

5.3.1 Solution en densité de courant

La figure 5.6 illustre la densité de courant $J_s(x, t)$ associée à la solution thermique de la figure 5.3 à divers instants. On observe une fois de plus que la solution du modèle RAPDi se superpose quasi parfaitement avec celle de la MEF 2-D, autant pour la longueur de transfert du courant que pour la zone résistive. On remarque toutefois que le modèle RAPDi omet un point à gauche de la zone résistance dans les figures 5.6b), c) et d) lorsque la densité de courant J_s est presque nulle. Cette erreur ponctuelle est observée vers $x \approx 2.6$ cm à la figure 5.6d). Ceci dit, la solution du modèle RAPDi reste exacte dans son ensemble, surtout pour un modèle pénalisant uniquement sur la température lors de son entraînement. On constate ainsi que, dès lors que la température $T(x, t)$ est exacte, l'équation de partage de courant (4.18) permet de reconstruire la solution J_s .

Le modèle de référence souffre quant à lui d'une solution thermique imprécise, ce qui entraîne une solution en densité de courant erronée. Pour les premiers instants où la température ne varie que d'une fraction de degré, la solution à la figure 5.6a semble exacte. Une variation de 0.5 K pour une température de 78 K n'a pas encore d'impact conséquent sur les propriétés électriques et thermiques. Cependant, on remarque que seulement 3 ms plus tard (figure 5.6b), la solution en densité de courant de la MDF 1-D commence à diverger de celle de la MEF 2-D. Le modèle de référence sous-estime la température, ce qui entraîne une proportion plus importante de courant dans la couche supraconductrice. Une fois que la zone normale est établie ($T > T_c$), tel qu'illustré aux figures c) et d), la solution en densité de courant est inexacte. Notons toutefois que la CTL semble être bien décrite, mais la largeur de la zone

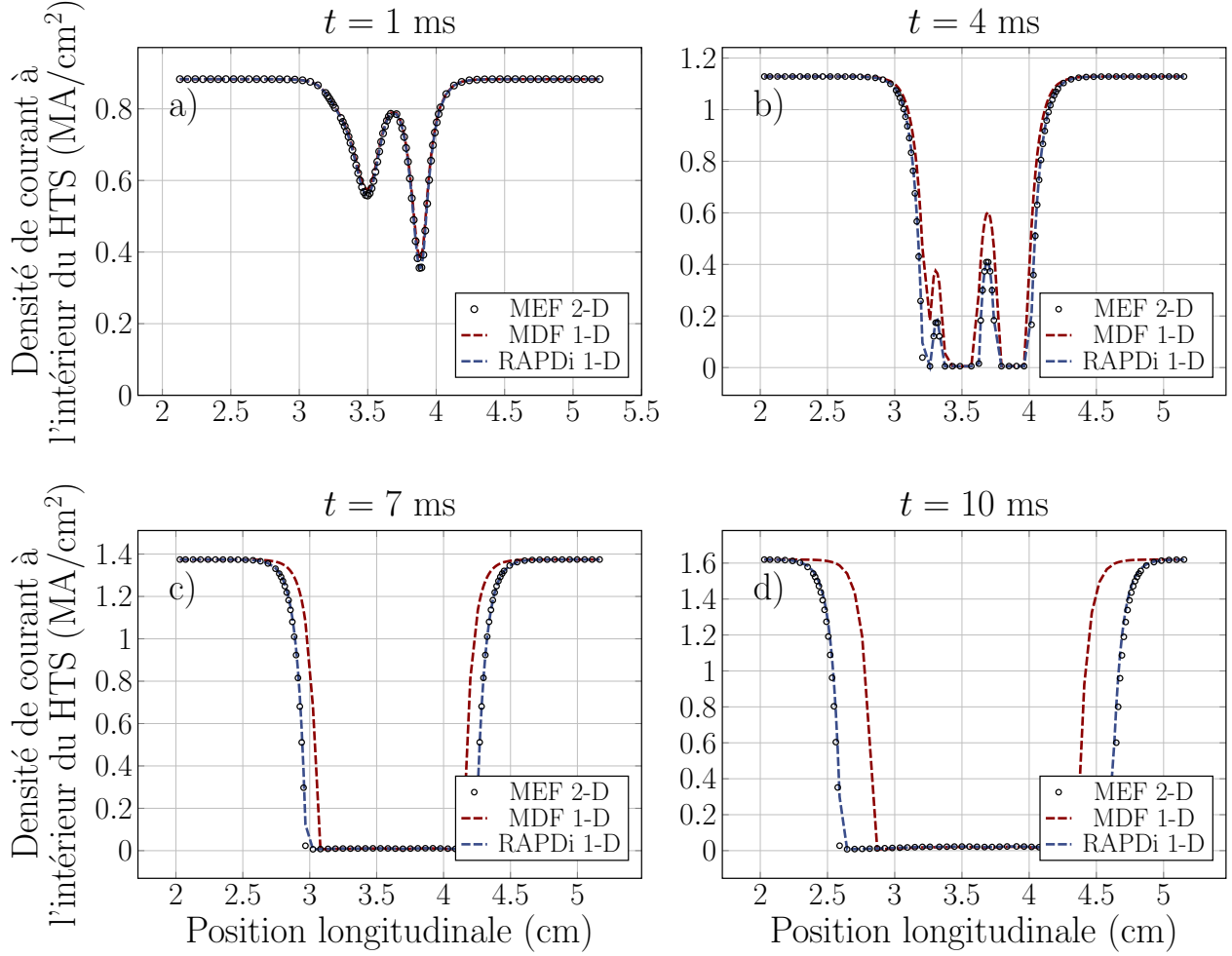


FIGURE 5.6 Densité de courant au milieu de la couche HTS d'un ruban uniforme aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms lors d'un régime dominé par les pertes Joule pour la MEF 2-D, la MDF 1-D (modèle de référence) et le modèle RAPDi. La résistance d'interface et l'épaisseur de la couche d'argent sont de $7.6 \mu\Omega \text{ cm}^2$ et $1.9 \mu\text{m}$, respectivement.

normale est décalée d'environ 0.35 cm de chaque côté vers le centre à $t = 10$ ms. Sur une échelle de 7 cm, il s'agit d'une erreur non négligeable.

5.3.2 Solution en potentiel électrique

La figure 5.7 illustre la solution en potentiel électrique à l'intérieur de la couche supraconductrice $V_s(x, t)$ associée à la solution thermique de la figure 5.3 à divers instants. On observe que le modèle RAPDi s'aligne une fois de plus avec la MEF 2-D, à l'inverse de la MDF 1-D. Comme le modèle de référence sous-estime la largeur de la zone résistive, la chute de tension de la MDF 1-D est toujours inférieure à celle de la MEF 2-D. On remarque également que la

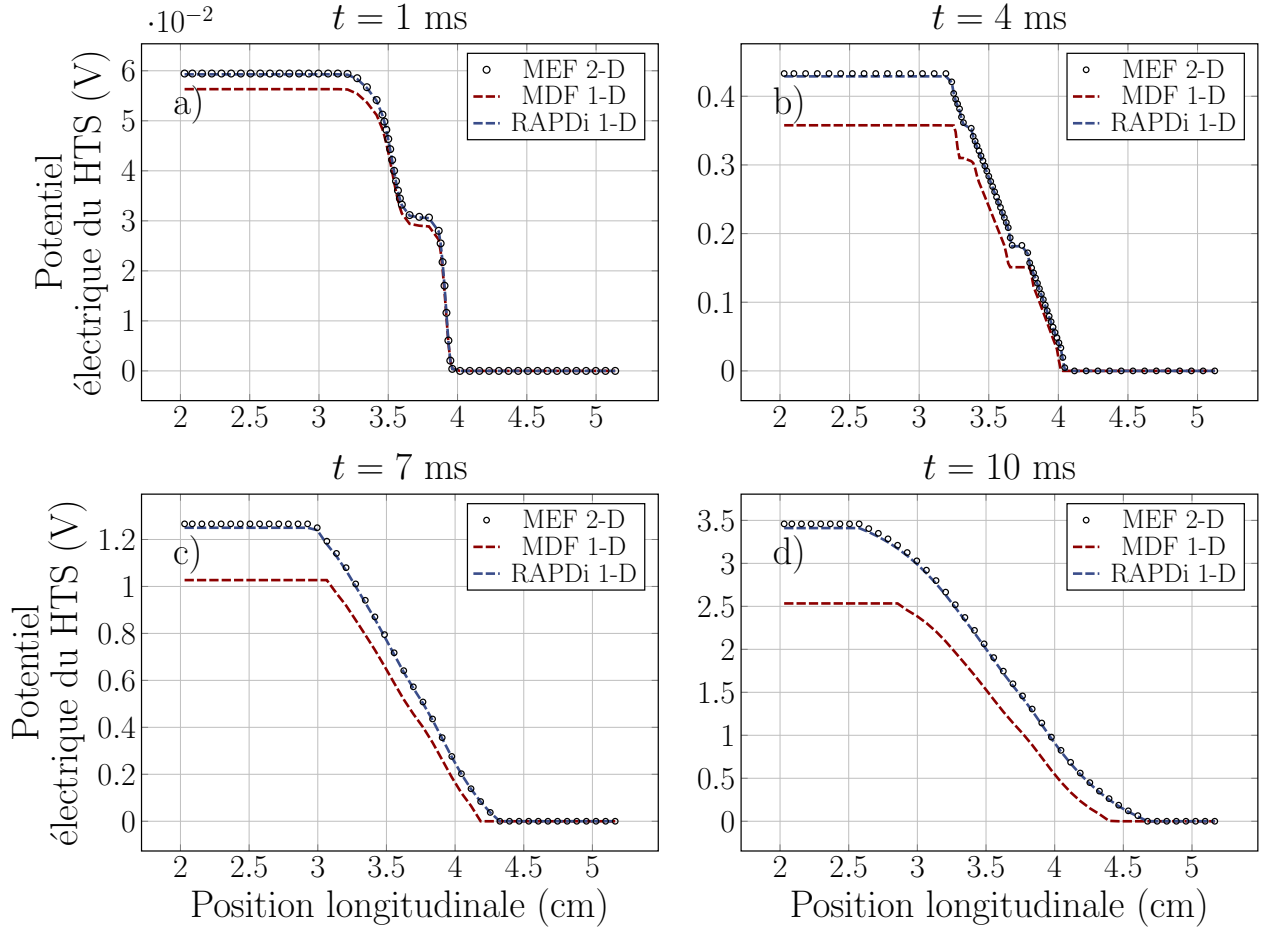


FIGURE 5.7 Potentiel électrique de la couche HTS d'un ruban uniforme aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms lors d'un régime dominé par les pertes Joule pour la MEF 2-D, la MDF 1-D (modèle de référence) et le modèle RAPDi. La résistance d'interface et l'épaisseur de la couche d'argent sont de $7.6 \mu\Omega \text{ cm}^2$ et $1.9 \mu\text{m}$, respectivement.

solution en potentiel électrique décrit mieux la qualité de la solution électrique, notamment lorsqu'on compare les figures 5.6a et 5.7a. La figure 5.6a semble indiquer que la solution de la MDF 1-D en densité de courant est identique à celle de la MEF 2-D, tandis que la figure 5.7a met en évidence la disparité entre leur solution V_s . Cette disparité entre les solutions obtenues peut être quantifiée à partir de l'erreur relative maximale

$$\Delta V_{rm}(t) = \max_x \left(\frac{|V_{MEF}(x, t) - V(x, t)|}{V_{MEF}(x, t)} \right), \quad (5.3)$$

où V_{MEF} est la solution par la MEF et V est la solution du modèle qu'on cherche à comparer. En ce sens, le tableau 5.6 présente l'erreur relative maximale des deux modèles 1-D à chaque

TABLEAU 5.6 Erreur relative maximale de la solution $V_s(x, t)$ pour un ruban uniforme à différents instants pour la MDF 1-D et le modèle RAPDi.

Temps	ΔV_{rm} MDF 1-D	ΔV_{rm} RAPDi 1-D
1 ms	5.26%	0.277%
4 ms	17.4%	0.951%
7 ms	18.9%	1.30%
10 ms	26.8%	1.41%

instant de la figure 5.7. On observe que l'erreur est croissante comme attendu. La MDF 1-D présente quant à elle une erreur relative maximale de 5% dès 1 ms, contrairement à 0.277% pour le modèle RAPDi. Cette même erreur atteint un seuil plafond de 26.8% à $t = 10$ ms, tandis que celle du modèle proposé est encore inférieure 1.5%. La solution en potentiel électrique V_s peut donc être obtenue seulement en pénalisant le problème thermique du modèle RAPDi lors de l'entraînement.

Le modèle RAPDi est donc un outil rapide et précis pouvant potentiellement être couplé à des logiciels de simulation de réseaux électriques, tel que EMTP-RV, afin de permettre la modélisation de points chauds dans un limiteur de courant HTS de plusieurs dizaines de mètres à faibles couts. Un tel couplage permettrait alors de modéliser les phénomènes microscopiques liés à la dynamique de quench des rubans HTS dans la modélisation de systèmes électriques à grande échelle, un problème qui n'a jamais pu être résolu à ce jour. La section qui suit donne un aperçu d'un tel cas.

5.4 Modélisation d'un ruban HTS uniforme de grande longueur

Le modèle NN-Moyen-3, entraîné sur des rubans d'une longueur $L = 7$ cm, est utilisé pour modéliser un ruban d'une longueur de $N_s \times L$ grâce à la propriété d'extension du domaine présentée à la section 4.4.5. Plus précisément, 20 simulations caractérisées par une rampe de courant fixe, une résistance d'interface $R_{inter} = 5 \mu\Omega \cdot \text{cm}^2$ ainsi qu'une épaisseur de stabilisateur $h_a = 2 \mu\text{m}$, sont concaténées spatialement afin d'obtenir artificiellement la solution de la MEF 2-D pour un ruban uniforme d'une longueur de 1.4 mètre. Les inhomogénéités en courant critique ont quant à elles été produites aléatoirement. Ceci dit, l'intervalle de valeurs de l'amplitude d'un défaut $\bar{A}_i$, initialement de $[0.55, 0.75]$ (tableau 4.2) a été légèrement élargi à $[0.45, 0.75]$. Cette modification permet de diversifier davantage la température maximale de chacune des simulations, puisque réduire la valeur de $\bar{A}_i$ implique un défaut moindre, et par conséquent, une génération de chaleur moins importante. Toutefois, comme le modèle RAPDi est entraîné selon le contexte thermique $(Q_{ag}, Q_s, Q_{inter}, Q_{conv})$, où la chaleur varie entre 0

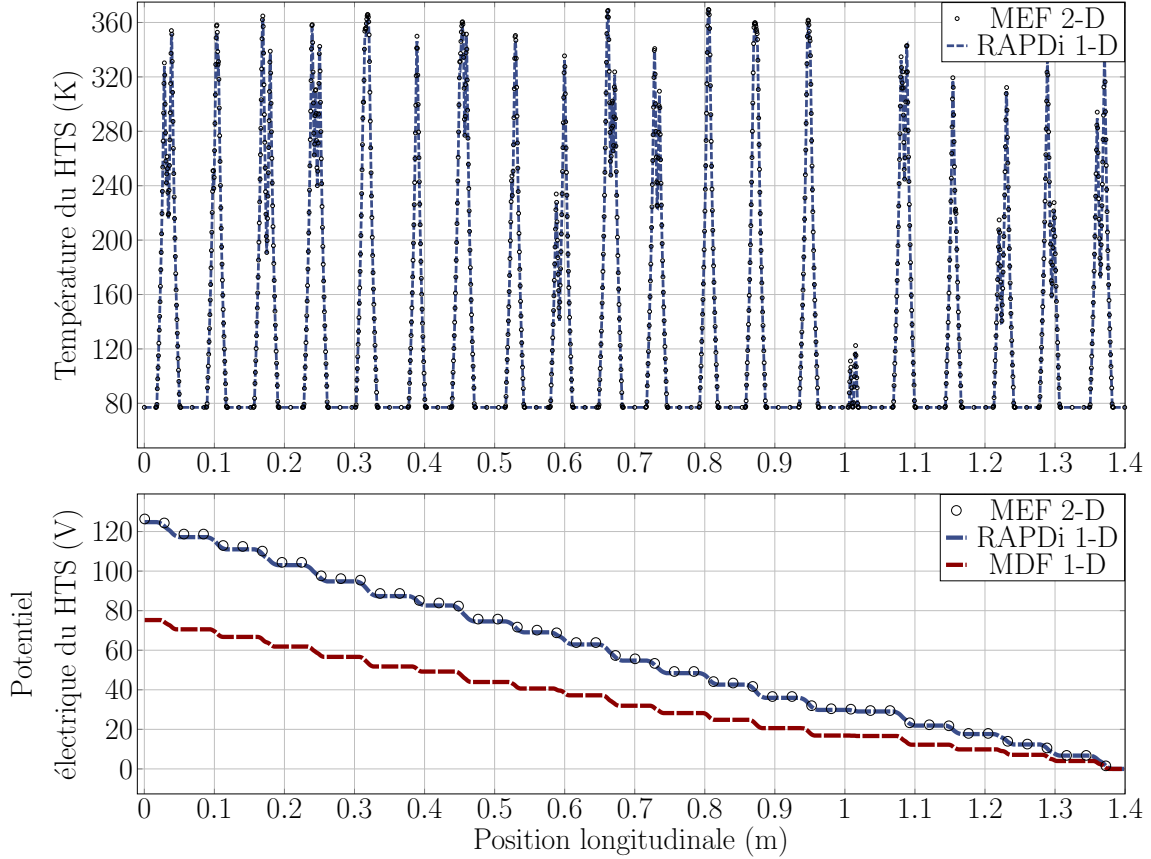


FIGURE 5.8 Solution électrothermique à l'intérieur de la couche supraconductrice d'un ruban HTS uniforme de 1.4 m à 10 ms où les pertes Joule sont dominantes. Les solutions en (a) température et (b) potentiel électrique obtenues par le modèle RAPDi sont comparées à celles de la MEF 2-D. La résistance d'interface est de $5 \mu\Omega \text{ cm}^2$ et l'épaisseur de stabilisateur est de $2 \mu\text{m}$. Ces valeurs n'ont jamais été vues durant l'entraînement du modèle.

et 10^7 W/m^2 , la chaleur observée une fois $\bar{A}_i$ réduit reste comprise dans l'intervalle (0 à 10^7 W/m^2) avec lequel le modèle a été entraîné.

Le modèle RAPDi prédit ainsi la solution du ruban de 1.4 mètre comme un seul ruban ayant n_D inhomogénéités selon la longueur. La solution électrothermique obtenue en température et potentiel électrique à $t = 10 \text{ ms}$ est présentée à la figure 5.8. Du côté de la solution en température, le modèle RAPDi 1-D semble concorder parfaitement avec la MEF 2-D. La solution en potentiel électrique $V_s(x, t)$ est quant à elle aussi précise, avec une erreur relative maximale de seulement 1.3%. À titre de référence, la MDF 1-D atteint une erreur relative maximale de 40%.

Du côté des performances temporelles, le temps de calcul du ruban de 1.4 mètre s'élève à 1.6 minute pour le modèle RAPDi, tandis que la MEF 2-D requiert 30h ($20 \times 1.5 \text{ h}$) pour

TABLEAU 5.7 Temps de calcul du modèle RAPDi 1-D pour différentes longueurs de ruban.

Longueur du ruban (m)	Temps (min)	Nombre de défauts n_D
1.4	1.6	74
2.8	3.7	148
14	25.5	740

cette même solution. Le modèle proposé est ainsi 1110 fois plus rapide que la méthode des éléments finis. Le temps de résolution pour différentes longueurs de rubans est également présenté au tableau 5.7. Notons que pour les rubans de plus de 1.4 mètre, la solution de la figure 5.8 a simplement été dupliquée comme on s'intéresse à la performance temporelle. Le modèle RAPDi permet ainsi de simuler des rubans jusqu'à 14 mètres de longueur, formés de centaines d'inhomogénéités, en 30 minutes. Toutefois, n'oublions pas que le facteur limitant en temps de calcul du modèle proposé demeure la résolution de l'équation (4.18) en raison de l'utilisation de méthodes numériques implicites.

CHAPITRE 6 MODÉLISATION DE RUBANS CFD D'ARCHITECTURE 3-D

Le principal inconvénient de la modélisation de rubans uniformes 2-D selon le modèle de Levin *et al.* réside dans l'incapacité de modéliser précisément le couplage thermique entre la couche supraconductrice et le substrat, tel que discuté au chapitre précédent. Cependant, le passage en géométrie 3-D introduit une complexité supplémentaire : le partage de courant n'a plus seulement lieu selon l'épaisseur, mais également selon la largeur. La modélisation de rubans CFD est notamment un cas où la distribution de courant devient non uniforme sur la section du ruban. Néanmoins, le modèle RAPDi est en mesure de modéliser le comportement électrothermique d'un ruban CFD malgré la nature tridimensionnelle des phénomènes étudiés.

Afin de démontrer la véracité de cette affirmation, on entame ce chapitre avec les phénomènes tridimensionnels à considérer pour les rubans CFD 3-D, les spécificités du jeu de données ainsi que l'introduction de nouvelles variables pour les modèles 1-D. Par la suite, l'entraînement et l'optimisation de l'architecture du modèle RAPDi sont effectués. Le chapitre se conclut avec la présentation des solutions électrothermiques d'un ruban CFD réalisées avec le modèle.

6.1 Du ruban CFD 3-D au ruban 1-D : les considérations

Les rubans CFD possèdent non seulement deux résistances d'interface distinctes, mais également des effets thermiques 3-D issus de la répartition non uniforme de la densité de courant J au sein du ruban. En ce sens, il est crucial de définir les solutions longitudinales $T(x, t)$ et $J(x, t)$ que le modèle RAPDi cherche à modéliser. Certes, le modèle RAPDi modélise le milieu de la couche supraconductrice (axe z), mais l'emplacement à privilégier selon la largeur (axe y) n'est pas nécessairement évident à déterminer. En conséquence, on définit deux emplacements, soit la droite $A - A'$ sur le bord du ruban, ainsi que la droite $B - B'$ au centre du ruban, tel qu'illustré à la figure 6.1. La droite $A - A'$ est centrée en $y = w_R/2$, où w_R est la largeur de l'interface résistive faible R_{inter} . Les solutions électrothermiques en ces deux emplacements sont alors comparées dans les prochaines sous-section afin de déterminer l'emplacement idéal.

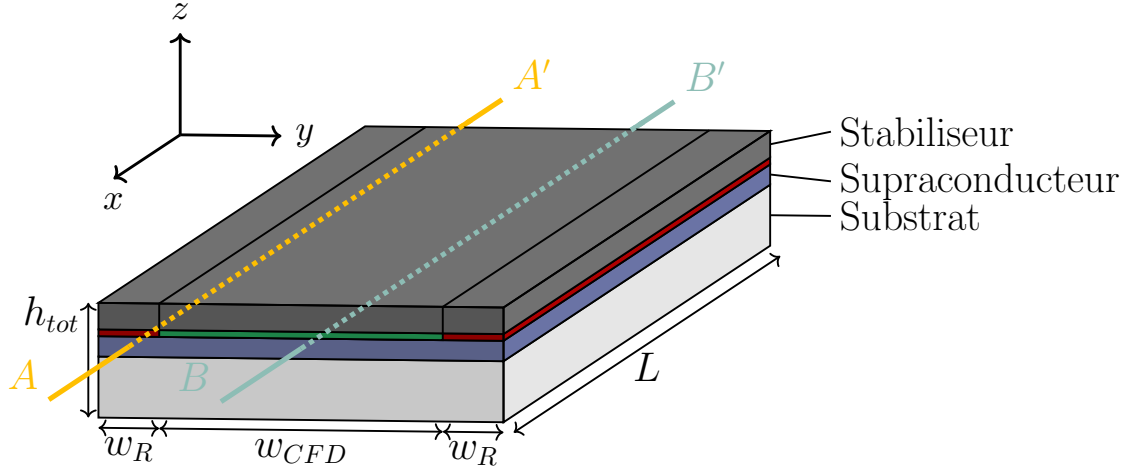


FIGURE 6.1 Schéma 3-D d'un ruban CFD : On considère les solutions longitudinales le long des lignes $A-A'$ et $B-B'$. La couche rouge représente la résistance d'interface R_{inter} (faible), tandis que la couche verte la résistance CFD R_{CFD} (élevée), et w_{CFD} est la largeur de la couche CFD. Les couches tampon ont été omises du schéma.

6.1.1 Distribution de température

La figure 6.2 montre la vue du dessus de l'évolution de la distribution de température dans un ruban CFD simulé par la MEF 3-D en présence d'inhomogénéités. Le plan de symétrie $x-z$ au centre du ruban permet par ailleurs de simuler des rubans de largeur $w/2$ sans perte de généralité. On observe également l'effet CFD à la figure 6.2a sous la forme d'une propagation non uniforme de la température, conséquence de la présence de la résistance d'interface CFD (R_{CFD}) au centre du ruban. L'architecture CFD favorise ainsi une propagation plus rapide de la zone normale vers les extrémités du ruban.

L'effet CFD n'est toutefois pas le seul effet thermique qui altère la distribution de température. Un phénomène de diffusion thermique est également observé, notamment lorsqu'on compare les points P_A et P_B de la figure 6.2b. Au point P_A , la variation de la température longitudinale ($\frac{\partial T}{\partial x}$) dans le régime transitoire est très abrupte, semblable au comportement d'un ruban uniforme (modèle de Levin *et al.*). La vitesse de propagation de la zone normale au point P_A est également plus élevée le long du ruban (axe x) que selon sa largeur (axe y), tel qu'illustré à la figure 6.2a à $t = 3$ ms. Ainsi, la chaleur générée au point P_A et dans son voisinage se diffuse sur la largeur du ruban avec un certain délai. Au point P_B , la chaleur provient non seulement de l'effet Joule local lorsque le courant transite entre la couche HTS et la couche stabilisatrice, mais aussi de la diffusion due à la température plus élevée sur le bord du ruban (point A et son voisinage). Le faible taux de variation de la température longitudinale

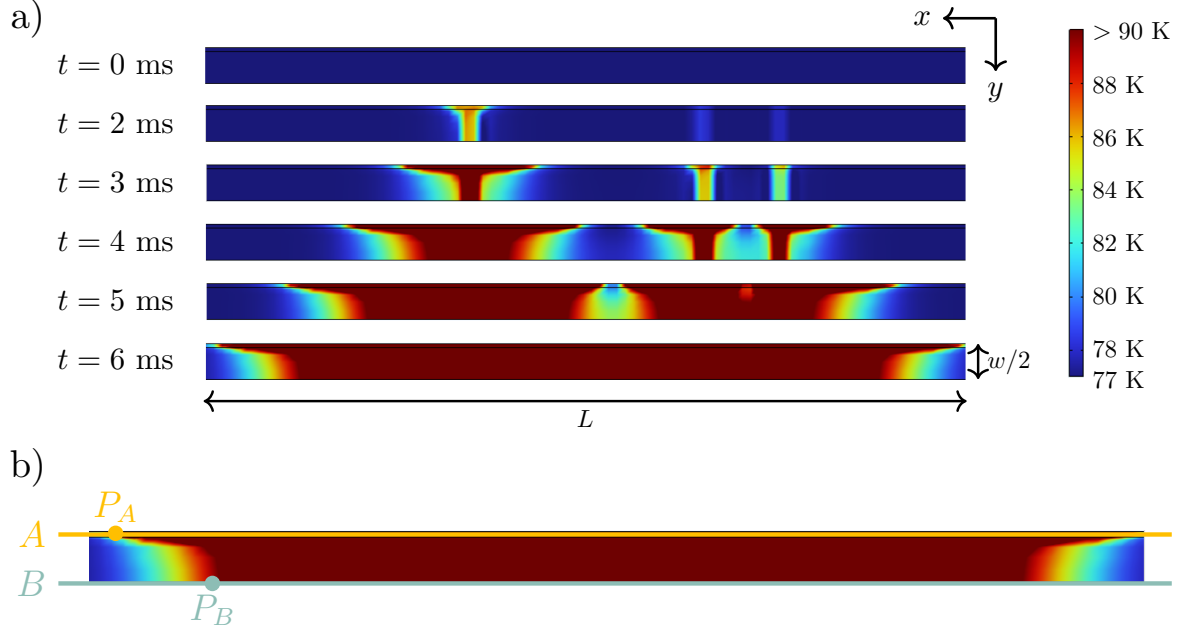


FIGURE 6.2 Vue de dessus du profil de température lors d'un quench pour un ruban CFD simulé par la MEF 3-D avec 3 défauts de tailles différentes sur sa longueur (a) à divers instants et (b) à l'instant final (6 ms).

au point P_B s'explique donc par la contribution de la diffusion thermique, contrairement au point P_A . La figure 6.3 illustre notamment le profil de température à $t = 6$ ms le long de la droite $A - A'$ ($T_{A-A'}$), le long de la droite $B - B'$ ($T_{B-B'}$) ainsi que le profil de température moyen $\bar{T}$ sur la section transverse du HTS définie comme suit :

$$\bar{T}(x) = \frac{1}{A_s} \int_{A_s} T(x, y, z) dy dz. \quad (6.1)$$

La figure 6.3 met en évidence la transition exponentielle entre l'état supraconducteur et l'état normal le long de la ligne $A - A'$ (transition de 77 K à 90 K près de P_A). Cette transition abrupte est caractéristique des rubans uniformes en raison de la résistance d'interface faible qui s'y retrouve, et donc de leur faible longueur de transfert (CTL). Cette transition exponentielle n'est toutefois pas observée le long de la ligne $B - B'$. En effet, la diffusion thermique rend la transition entre l'état supraconducteur et l'état normal quasi-linéaire.

Ceci dit, l'effet de diffusion thermique sur la largeur du ruban (axe y) ne s'exprime pas naturellement par un modèle 1-D. Cette dissipation ajoute une complexité thermique supplémentaire, s'ajoutant aux effets 3-D déjà existants. En ce sens, afin de favoriser l'apprentissage du modèle RAPDi ainsi que pour maximiser l'exactitude du modèle de Levin *et al.*, il est

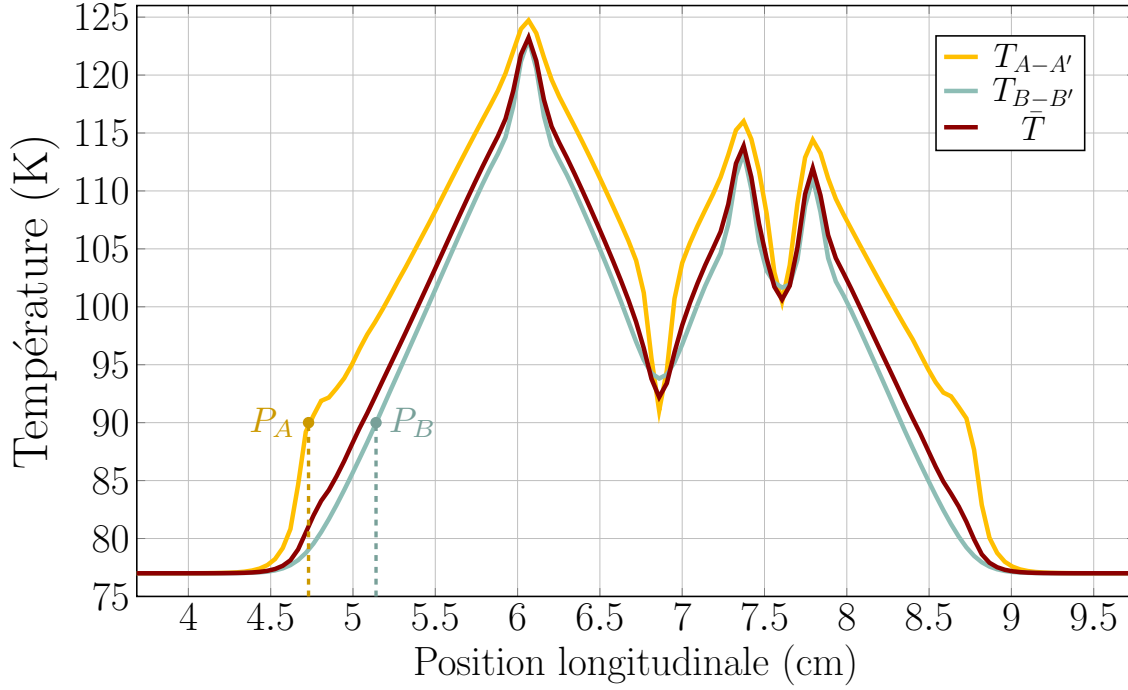


FIGURE 6.3 Température de la couche HTS d'un ruban CFD 3-D le long des lignes $A - A'$ ($T_{A-A'}$) et $B - B'$ ($T_{B-B'}$) et la température moyenne $\bar{T}$. Les points P_A et P_B de la figure 6.2 sont indiqués.

préférable d'éviter de modéliser la solution le long de la ligne $B - B'$.

6.1.2 Résistance d'interface effective

Le ruban CFD possède deux résistances d'interface ayant chacune leur rôle : l'une est majoritairement responsable de la CTL (R_{CFD}) et la seconde est responsable de la génération de chaleur (R_{inter}). En ce sens, l'idée intuitive pour modéliser un ruban CFD selon un modèle 1-D est de définir une résistance d'interface effective $R_{inter,T}$ pour le problème thermique, et une seconde résistance d'interface effective $R_{inter,J}$ pour le problème électrique pour reproduire la bonne longueur de transfert. Toutefois, cette approche est limitée par le modèle 1-D utilisé. Plus précisément, étant donné que le transfert de courant se fait dans un environnement tridimensionnel, rien ne garantit que le modèle 1-D est capable d'exprimer la même « famille » de solutions en densité de courant $J_s(x)$ que celle obtenue par la MEF 3-D. Le modèle de Levin *et al.* n'y échappe pas : l'équation du partage de courant (4.18) permet uniquement de modéliser des transferts de courant de nature exponentielle (rubans uniformes).

Ceci dit, pour la suite de ce chapitre, on définit $R_{inter,J}$ comme la résistance d'interface qui minimise l'erreur de reconstruction de la solution $J_s(x)$ entre le modèle de référence et la

MEF 3-D au temps final. On exprime alors mathématiquement $R_{inter,J}$ tel que

$$R_{inter,J} = \underset{R_{inter}}{\operatorname{argmin}} \left\| \mathbf{J}_s^{Levin} - \mathbf{J}_s^{MEF} \right\|_2, \quad (6.2)$$

où $\mathbf{J}_s^{Levin}$ et $\mathbf{J}_s^{MEF}$ dénotent respectivement la solution discrétisée en densité de courant à l'intérieur de la couche supraconductrice selon le modèle de Levin *et al.* et selon la MEF 3-D. Notons également que la solution $\mathbf{J}_s^{Levin}$ est calculée à partir de la solution thermique $T(x, t)$ de la MEF 3-D, donc seule la variable R_{inter} de l'équation (4.18) est optimisée.

6.1.3 Longueur de transfert

La longueur de transfert en densité de courant le long des lignes $A - A'$, $B - B'$ ainsi que la densité de courant moyenne $\bar{J}_s(x)$ sont présentées à la figure 6.4. Notons que chacune des solutions obtenues par le modèle de Levin *et al.* a été calculée à partir de la résistance d'interface effective $R_{inter,J}$. De plus, $\bar{J}_s(x)$ est définie telle que

$$\bar{J}_s(x) = \frac{1}{A_s} \int_{A_s} J_s(x, y, z) dy dz. \quad (6.3)$$

La figure 6.4a illustre une superposition entre la MDF 1-D et la MEF 3-D pour la grande majorité de la longueur de transfert. Cela signifie donc que les phénomènes électriques qui ont lieu le long de $A - A'$ sont approximativement de la même nature que ceux modélisés par le modèle de Levin *et al.*, soit une longueur de transfert de nature exponentielle. Toutefois, la figure 6.4b reflète un tout autre cas de figure : la MDF 1-D n'arrive pas à modéliser la longueur de transfert au milieu du ruban, et ce, malgré l'optimisation de la valeur de $R_{inter,J}$. En fait, on constate que non seulement la longueur de transfert est plus longue le long de la ligne $B - B'$ que le long de la ligne $A - A'$, mais que la forme de celle-ci est également différente. La CTL à la figure 6.4a est bien plus abrupte, tandis que celle de la figure 6.4b illustre une transition plus douce, s'apparentant à une fonction sigmoïde. En ce sens, la MDF 1-D n'est pas capable de reconstruire le profil $J_s(x)$ correctement, puisque le modèle de Levin *et al.* ne décrit pas la même famille de fonctions que la MEF 3-D. La figure 6.4c illustre quant à elle un cas amélioré de la figure 6.4b puisque celle-ci représente une moyenne sur la surface transverse. En ce sens, la densité de courant $\bar{J}_s(x)$ pondère autant la solution sur le bord du ruban (figure 6.4a) que la solution en son centre (figure 6.4b). Par ailleurs, notons que les erreurs commises par la MDF 1-D à la figure 6.4 sont seulement le produit d'une seule prédiction. Lorsqu'on calcule itérativement la solution jusqu'à plus de 500 instances (N_t), comme dans le chapitre précédent, l'accumulation d'erreurs peut alors devenir considérable.

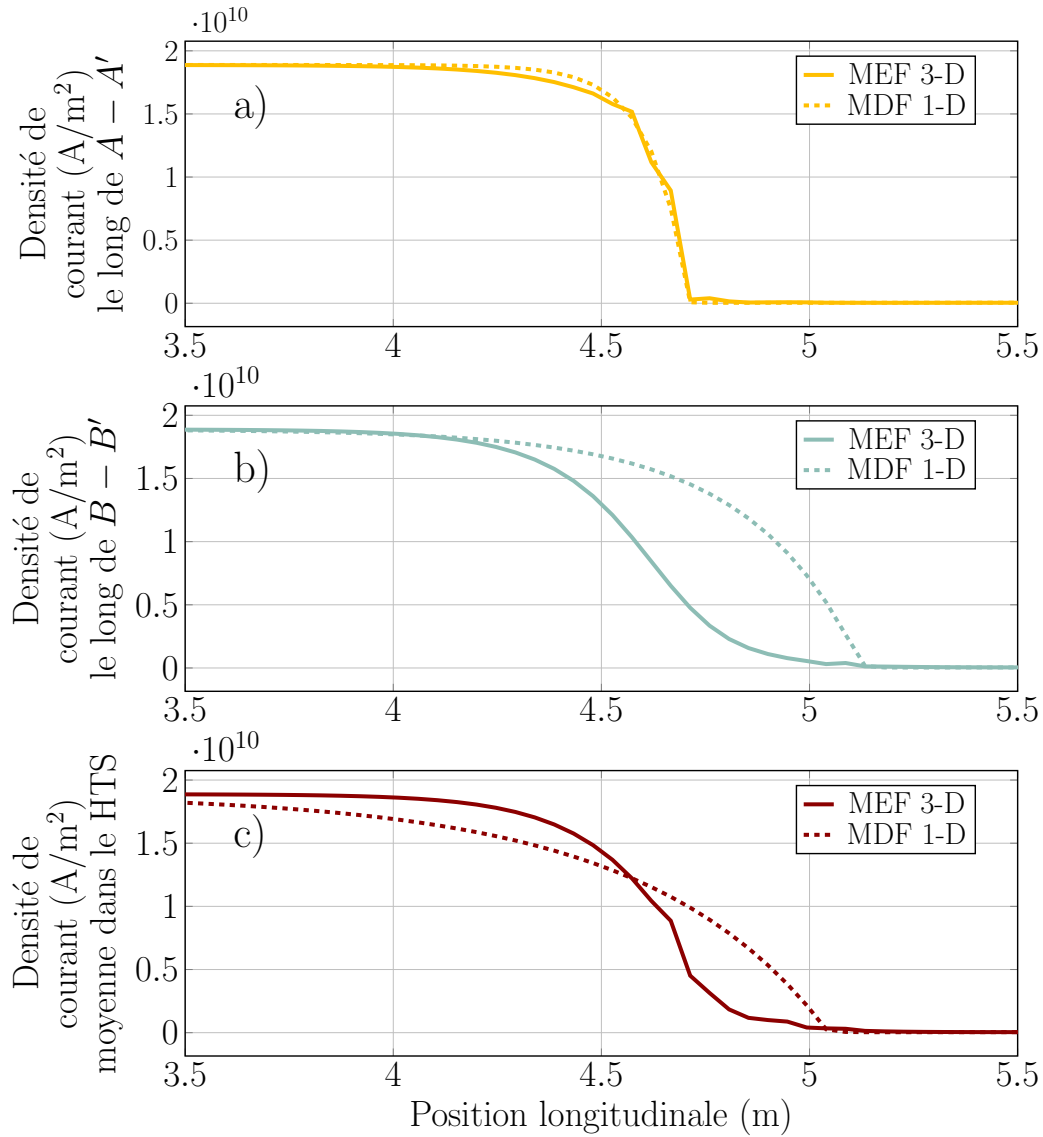


FIGURE 6.4 Longueurs de transfert en densité de courant d'un ruban CFD 3-D (a) le long de la droite $A - A'$, (b) le long de la droite $B - B'$ et (c) sur la section transverse. La résistance d'interface effective $R_{inter,J}$ est de $6.2 \mu\Omega.cm^2$, $230 \mu\Omega.cm^2$ et $150 \mu\Omega.cm^2$ pour $A - A'$, $B - B'$ et la section transverse respectivement.

Ceci dit, les figures 6.3 et 6.4 démontrent un résultat insoupçonné jusqu'à aujourd'hui : les bords d'un ruban CFD se comportent *localement* comme un ruban uniforme. Le mot clé ici est « localement », car bien évidemment, un ruban uniforme 2-D n'est pas en mesure de reproduire la dynamique d'un ruban CFD. Il est nécessaire de prendre en considération la

chaleur générée par la déviation de J le long de la largeur du ruban ainsi que la dissipation de chaleur vers le milieu du ruban, soit deux phénomènes purement 3-D.

À la lumière de ces résultats, le modèle RAPDi sera donc utilisé pour modéliser la solution électrothermique du ruban CFD le long de $A - A'$ pour le reste de ce mémoire. Étant donné que le bord du ruban CFD peut être modélisé approximativement comme un ruban uniforme 1-D et que le modèle RAPDi est basé sur le modèle de Levin *et al.*, il est tout à notre avantage de procéder ainsi afin de maximiser la précision du modèle. Dans un certain sens, bien que la solution électrothermique sur le bord du ruban et celle au milieu du ruban soient différentes, les deux emplacements fournissent des informations similaires quant à l'état thermique du ruban (état de quench), et l'état électrique (V_s et CTL).

6.2 Génération de données

L'ensemble des simulations de ce chapitre est généré à l'aide du modèle électrothermique par éléments finis défini à la section 4.2.1. Le modèle par éléments finis n'est toutefois plus utilisé afin de résoudre des problèmes 2-D, mais plutôt des problèmes 3-D où la géométrie étudiée est celle d'un ruban CFD enrobé (architecture présentée à la figure 2.7b). Les rubans CFD enrobés sont notamment privilégiés aux rubans CFD classiques puisque ceux-ci permettent d'accélérer davantage la vitesse de propagation de la zone normale (NZPV) via l'ajout de masse thermique. Ceci dit, certaines modifications doivent être effectuées en raison du changement de géométrie et du passage en trois dimensions. Les prochaines sous-sections visent à détailler ces changements.

6.2.1 Paramètres géométriques et inhomogénéités aléatoires

La principale caractéristique des rubans CFD est leur vitesse de propagation de la zone normale (NZPV) élevée contrairement aux rubans uniformes. Il est en conséquence nécessaire de modifier les valeurs de certains paramètres géométriques si l'on veut s'assurer de maintenir des zones tampons aux extrémités du ruban CFD. Les paramètres géométriques considérés pour un ruban CFD enrobé sont donc présentés au tableau 6.1. On remarque tout d'abord que la longueur du ruban CFD enrobé a été doublée ($7 \rightarrow 14$ cm) comparativement aux rubans uniformes. Les paramètres h_a , $h_{a,bot}$ et $h_{a,side}$ servent quant à eux à quantifier l'épaisseur du revêtement entourant le ruban, fixée ici à 2 μm .

Notons toutefois que la résistance d'interface R_{inter} ainsi que l'épaisseur de la couche de stabilisation h_a indiquée au tableau 6.1 sont maintenues fixes, contrairement au cas d'un ruban uniforme (voir tableau 4.1). Ce choix vise à simplifier la tâche d'apprentissage, puisque

TABLEAU 6.1 Paramètres géométriques de rubans CFD enrobés.

Paramètres	Valeur numérique
Longueur (L)	14 cm
Largeur (w)	4 mm
Épaisseur de substrat (h_{hast})	50 μm
Épaisseur couche tampon (h_{MgO})	150 nm
Épaisseur couche supraconductrice (h_s)	2 μm
Résistance d'interface (R_{inter})	1 $\mu\Omega \text{ cm}^2$
Résistance CFD (R_{CFD})	10 000 $\mu\Omega \text{ cm}^2$
Épaisseur couche stabilisatrice supérieure (h_a)	2 μm
Épaisseur couche stabilisatrice inférieure ($h_{a,bot}$)	2 μm
Épaisseur couche stabilisatrice latérale ($h_{a,side}$)	2 μm

la modélisation de phénomènes 3-D à partir d'un modèle 1-D est une tâche déjà complexe en soi. Du côté des inhomogénéités de la densité de courant critique, celles-ci sont variées aléatoirement de la même façon que précédemment. Les propriétés géométriques restantes de la couche HTS sont également inchangées.

Pour les paramètres de résolution, le modèle électrothermique 3-D par éléments finis résout la dynamique du quench dans un ruban CFD avec un maillage $\Delta x = 0.47 \text{ mm}$ ($N_x = 301$) et un pas de temps $\Delta t = 33 \text{ } \mu\text{s}$ ($N_t = 301$). Un maillage plus grossier que pour un ruban régulier peut ici être utilisé, étant donné que la CTL est généralement plus élevée. Toutefois, le maillage selon x ne doit pas être réduit outre mesure, car $R_{inter,J}$ est identifié à partir d'une courbe telle qu'à la figure 6.4, qui présente une CTL « locale » plus petite que la CTL « globale » du ruban CFD. Un maillage irrégulier composé de 14 points (N_y) est utilisé en largeur afin de densifier les éléments près des bords du ruban. Du côté de l'épaisseur, on retrouve également un maillage irrégulier formé de 6 éléments ($N_z = 7$). Une simulation d'un ruban CFD enrobé de 14 cm (48 000 degrés de liberté) dure en moyenne 3.7 heures.

6.2.2 Rampes de courant aléatoires

Les rampes de courant représentant un régime dominé par le refroidissement convectif sont écartées lors de la génération des données 3-D. La solution électrothermique associée à ces rampes varie peu sur une échelle de 10 ms, en plus de posséder une faible accumulation d'erreur. En ce sens, comme le chapitre précédent a démontré qu'il était possible d'apprendre les deux régimes simultanément, le présent chapitre se concentre uniquement sur le régime dominé par les pertes Joule, là où l'accumulation d'erreur est la plus importante.

Ceci dit, la génération de rampes aléatoires a un impact important sur le modèle 1-D utilisé.

Plus précisément, la dynamique 3-D du quench dans un ruban CFD dépend du courant appliqué $I_{app}(t)$ en raison de l'augmentation de la densité de courant sur les bords du ruban. En ce sens, si on modélise cette même dynamique à partir du modèle de Levin *et al.* (modèle 1-D) à l'aide d'une résistance d'interface effective $R_{inter,J}$, alors $R_{inter,J}$ acquiert une dépendance au courant $I_{app}(t)$. Cette dépendance survient en raison du fait que le paramètre $R_{inter,J}$ cherche à modéliser la dynamique du ruban CFD, une dynamique dépendante de $I_{app}(t)$. Il en va de même pour $R_{inter,T}$. Toutefois, cette dépendance entre les résistances d'interface effectives et $I_{app}(t)$ est seulement observée au niveau du modèle 1-D. Les paramètres R_{inter} et R_{CFD} sont indépendants du courant dans le problème d'éléments finis 3-D.

Il advient ainsi que si la rampe de courant est aléatoire, alors le modèle de Levin *et al.* doit nécessairement utiliser des résistances d'interface effectives $R_{inter,J}$ et $R_{inter,T}$ différentes pour chacune de ces rampes. Néanmoins, comme l'effet dominant lors du quench est l'effet CFD, soit un effet électrique, on peut concentrer notre analyse uniquement sur $R_{inter,J}$. Le problème revient à trouver une façon d'exprimer $R_{inter,T}$ à partir de $R_{inter,J}$ et des autres informations du problème. En ce sens, la figure 6.5 illustre la relation statistique entre $R_{inter,J}$ et $I_{app}(t)$ au temps final (s'applique uniquement aux modèles 1-D) sous la forme d'un nuage de points pour des dizaines de simulations. Rappelons également que $R_{inter,J}$ a été déterminé à l'aide de l'équation (6.2) pour chacune des simulations. On aperçoit une légère corrélation

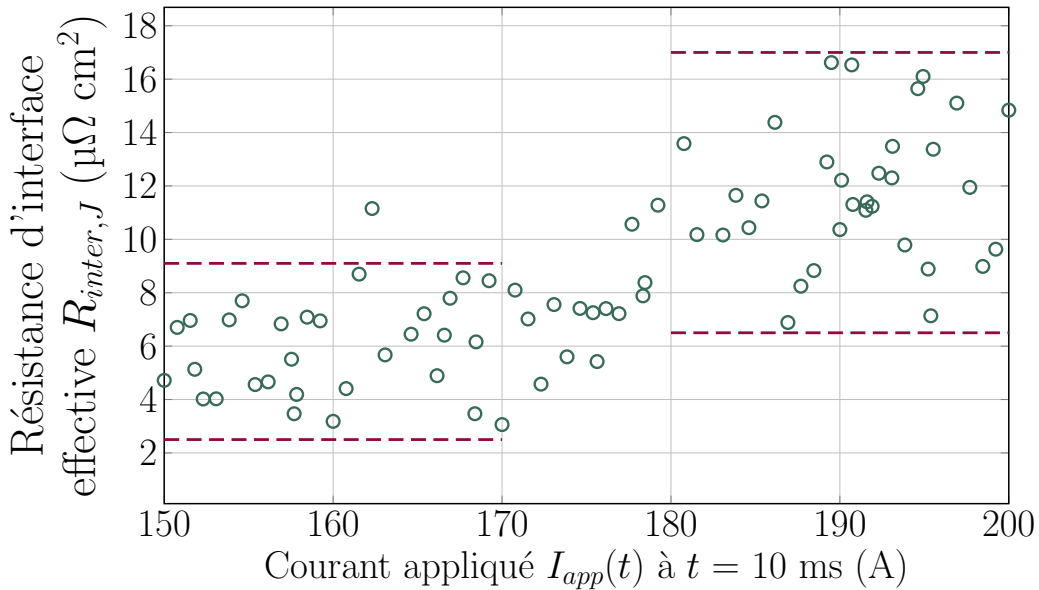


FIGURE 6.5 Nuage de point : Correspondance entre la résistance d'interface effective $R_{inter,J}$ et le courant appliqué $I_{app}(t)$ à l'instant final $t = 10$ ms pour des dizaines de simulations. Les intervalles rouges déterminent deux zones de valeurs de $R_{inter,J}$ distinctes.

entre l'intensité du courant et la résistance d'interface effective du problème électrique. On remarque toutefois une certaine transition des valeurs de $R_{inter,J}$ vers $I = 175$ A. Plus exactement, on semble observer une tendance telle que $R_{inter,J} < 9 \mu\Omega \text{ cm}^2$ lorsque $I < 170$ A et $R_{inter,J} > 7 \mu\Omega \text{ cm}^2$ pour $I > 180$. Dans chacun de ces intervalles, I_{app} ne semble pas avoir de pouvoir prédictif significatif. Ceci dit, il s'agit d'une tendance empirique méritant plus d'approfondissement quant à sa nature. L'hypothèse actuelle est que la figure 6.5 reflète indirectement la relation entre la longueur de transfert (CTL) et le courant appliqué I_{app} pour un ruban CFD, ce qui implique que plus le courant est élevé, plus la CTL est grande. Le lien entre $R_{inter,J}$ et la CTL vient de la définition même de l'équation (6.2) : $R_{inter,J}$ est optimisée de façon à reconstruire la solution J_s , et de ce fait, la longueur de transfert par la même occasion.

6.3 Modifications des modèles 1-D

En vue de modéliser les rubans CFD le long de la ligne $A - A'$, les modèles 1-D nécessitent d'incorporer les résistances d'interface effectives $R_{inter,J}$ et $R_{inter,T}$ au sein de leur architecture. Les sous-sections suivantes ont pour but de détailler comment le modèle de référence ainsi que le modèle RAPDi parviennent à incorporer ces paramètres.

6.3.1 Modèle de référence

Il est possible pour le modèle de Levin *et al.* d'approximer la résistance d'interface effective $R_{inter,J}$ à condition que la résistance CFD soit largement supérieure à la résistance faible ($R_{CFD} \gg R_{inter}$). Si tel est le cas, alors cette dernière s'exprime ainsi [24] :

$$R_{inter,J} \approx \frac{R_{inter}}{1 - f_{CFD}}, \quad (6.4)$$

où $f_{CFD} = \frac{w_{CFD}}{2w_R + w_{CFD}}$, et w_{CFD} est la largeur de la couche CFD. Néanmoins, cette approximation de $R_{inter,J}$ reste indépendante du courant appliqué, ce qui ne reflète pas les résultats empiriques observés à la figure 6.5 dans le cadre de modèles 1-D. Par ailleurs, rien n'assure que la résistance d'interface effective trouvée par l'équation (6.4) permette de reconstruire correctement la solution en densité de courant $J_s(x, t)$ selon $A - A'$.

Il ne faut cependant pas perdre de vue que le modèle de Levin *et al.* est un modèle de référence, c'est-à-dire un modèle décrivant les performances minimales que le modèle RAPDi doit obtenir. Par conséquent, nous ne sommes pas obligés de calculer $R_{inter,J}$ à partir de l'équation (6.4) si l'on veut maximiser la précision du modèle de référence. Dans cette optique, le modèle de référence pré-calcule plutôt $R_{inter,J}$ à partir de l'équation (6.2) pour chaque

simulation du jeu de données. La résistance d'interface effective $R_{inter,J}$ devient alors un paramètre qu'on fournit en entrée du modèle, remplaçant le paramètre géométrique R_{inter} dans le problème électrique.

Du côté du problème thermique, nous utiliserons simplement $R_{inter,T} = R_{inter,J}$ afin de rester cohérent avec un modèle classique (ruban uniforme). Il est délicat d'optimiser $R_{inter,T}$ puisque cette dernière peut non seulement varier en fonction du courant appliqué, mais également en fonction de l'état (supraconducteur/normal) dans lequel le ruban se trouve. Par ailleurs, on soupçonne que l'erreur du modèle 1-D sera dominée par l'erreur associée à la CTL en raison de la dominance des effets électriques tridimensionnels. L'équation thermique du modèle de référence (équation (4.14)) n'a alors pas nécessairement besoin d'être modifiée, puisque la valeur $R_{inter,T}$ n'a pas autant d'impact que son homologue $R_{inter,J}$. Enfin, notons que les surfaces latérales d'argent ne sont pas considérées lors de l'homogénéisation puisque cette quantité de masse thermique est négligeable ($w \gg h_{a,side} = 2 \mu\text{m}$).

6.3.2 Modèle RAPDi

Le modèle RAPDi ne calculera pas la résistance d'interface effective $R_{inter,J}$ via l'équation (6.4). Le modèle proposé se base plutôt sur une hypothèse forte : $R_{inter,J}$ est supposé connu à l'avance, comme c'est le cas avec le modèle de référence. Le modèle RAPDi traite alors la modélisation d'un ruban CFD comme un ruban uniforme ayant une résistance d'interface variable en fonction du courant appliqué au temps final $I_{app}(t_f)$. Une telle modélisation est possible, tel que montré au chapitre précédent lorsque le modèle RAPDi était entraîné pour différentes résistances d'interface au sein du même jeu de données. Par ailleurs, notons que l'introduction de cette hypothèse n'est pas tout-à-fait une limitation majeure. Il est possible de récolter plusieurs données afin de cartographier la relation entre le courant appliqué I_{app} et la résistance d'interface effective $R_{inter,J}$ sous la forme d'une fonction $f(I_{app})$. Cependant, l'estimation de la fonction $f(I_{app})$ n'est pas abordée dans ce mémoire. Cette tâche est laissée à des travaux ultérieurs.

Du côté thermique, le modèle RAPDi utilise simplement $R_{inter,J}$ comme interface résistive dans l'équation thermique (4.30). Toutefois, si on simplifie le terme de chaleur de la résistance d'interface (équation (4.32)) tel que

$$Q_{inter,i}^n d_\theta = h_a^2 d_\theta R_{inter,J} \left(\frac{\partial J_a(x_i, t_n)}{\partial x} \right)^2 = h_a^2 R_{inter,T} \left(\frac{\partial J_a(x_i, t_n)}{\partial x} \right)^2, \quad (6.5)$$

nous sommes alors capables de définir l'identité $R_{inter,T} = d_\theta R_{inter,J}$, où d_θ est un des préfacteur du modèle d'apprentissage profond *feed forward* présenté à la figure 4.11 (voir aussi

équation (4.32)). Non seulement le modèle RAPDi trouve une résistance d'interface effective $R_{inter,T}$ pour le problème thermique qui minimise l'erreur sur le jeu de données, mais celle-ci s'adapte également en fonction du contexte thermique en raison de la dépendance du préfacteur d_θ avec les variables T , Q_a , Q_s , Q_{inter} et Q_{conv} (figure 4.11). De l'autre côté, la résistance d'interface $R_{inter,J}$ du problème électrique est indépendante du contexte thermique. Cette dernière est donc maintenue fixe pour une simulation donnée. En ce sens, comme la résistance d'interface $R_{inter,J}$ est fixée lors de l'entraînement, la résistance d'interface effective du problème thermique $R_{inter,T}$ est apprise de manière à compenser l'erreur qu'engendre $R_{inter,J}$ en premier lieu.

6.4 Architecture et performances du modèle RAPDi

Les données recueillies pour l'entraînement du modèle RAPDi sont constituées de 82 simulations électrothermiques 3-D résolues par la méthode des éléments finis sur une géométrie de ruban CFD. Les simulations sont réparties en ensembles d'entraînement (62 simulations), de validation (1 simulation), et de test (19 simulations).

Le modèle RAPDi résout quant à lui la dynamique du quench le long de la ligne $A - A'$ avec la même discrétisation que la MEF 3-D, c'est-à-dire selon un maillage régulier de $\Delta x = 0.47$ mm ($N_x = 301$) et un pas de temps $\Delta t = 33$ μ m ($N_t = 301$). Du côté des hyperparamètres fixes lors de l'entraînement, on retrouve les fonctions d'activation (ReLU), la taille de mini-lot $\mathcal{B}$ (32), le nombre d'époques (200) et le taux d'apprentissage ($7.5 \cdot 10^{-5}$). Tous les paramètres du modèle RAPDi sont une fois de plus optimisés à l'aide de la méthode ADAM.

Ceci dit, on retrouve deux distinctions importantes entre le jeu de données des rubans CFD et celui des rubans uniformes (chapitre précédent). D'une part, le jeu de données des rubans CFD est moins abondant que celui des rubans uniformes. Non seulement les simulations sont moins nombreuses, mais le nombre de séquences vues durant l'entraînement l'est également ($N_t = 301$ pour les rubans CFD versus $N_t = 501$ pour les rubans uniformes). D'autre part, l'entraînement du modèle RAPDi est plus susceptible au surapprentissage, car le modèle de Levin *et al.* est imprécis lors de la modélisation de rubans CFD. Le modèle d'apprentissage profond intégré au modèle RAPDi doit alors compenser une erreur plus importante à partir des préfacteurs a_θ à f_θ , ce qui complexifie le processus d'optimisation de ses paramètres θ .

En ce sens, les prochaines sous-sections établissent l'espace de modèles RAPDi étudiés ainsi que les performances de chacun des modèles considérés, afin d'identifier le modèle le plus adéquat pour le jeu de données disponible. Cette section se conclut alors avec les performances temporelles du modèle RAPDi retenu.

TABLEAU 6.2 Modèles RAPDi étudiés à base de réseau *feed forward*.

Nom du modèle	Architecture	Dropout p	Taille fenêtre glissante
NN-Profond-1	[18,18,18,18,6]	0.075	3
NN-Profond-2	[30,30,30,30,30,6]	0.075	5
NN-Profond-3	[42,42,42,42,42,42,6]	0.075	7
NN-Petit-1	[18,75,75,6]	0.05	3
NN-Petit-2	[30,150,150,6]	0.05	5
NN-Moyen-1	[30,200,200,200,6]	0.075	5
NN-Moyen-2	[30,250,100,250,6]	0.075	5
NN-Moyen-3	[30,250,250,250,250,6]	0.075	5
NN-Gros-1	[30,250,250,250,250,250,6]	0.1	5
NN-Gros-2	[54,350,350,350,350,350,6]	0.15	9
NN-Gros-3	[150,400,400,400,400,400,400,6]	0.2	25
NN-Gros-4	[186,800,800,800,800,800,800,800,6]	0.2	31

TABLEAU 6.3 Modèles RAPDi étudiés à base de modèles d'APA.

Nom du modèle	N_B	Architecture FF	Dropout p	Taille fenêtre glissante
APA-Profond-1	6	[18,18]	0.05	3
APA-Profond-2	8	[30,30]	0.075	5
APA-Profond-3	12	[30,18,30]	0.075	5
APA-Petit-1	1	[150,150]	0.05	5
APA-Petit-2	1	[200,200]	0.075	5
APA-Moyen-1	1	[200,200,200]	0.075	5
APA-Moyen-2	2	[200,200]	0.075	5
APA-Moyen-3	3	[200,200]	0.075	5
APA-Gros-1	4	[350,350]	0.15	9
APA-Gros-2	4	[350,350,350]	0.15	15
APA-Gros-3	4	[400,400,400]	0.2	25
APA-Gros-4	5	[600,600,600]	0.2	31

6.4.1 Espace de modèles étudiés

L'espace de modèles RAPDi étudiés est défini par le type et par la taille du modèle d'apprentissage profond formant le modèle RAPDi, par la probabilité de *dropout* p , ainsi que par la taille de la fenêtre glissante utilisée, tel que présenté à la section 5.1.1. En ce sens, 10 architectures de modèle RAPDi à base de réseau *feed forward*/réseau de neurones (architecture présentée à la figure 4.11) ont été étudiées. Ces diverses architectures sont présentées au tableau 6.2. La notation [18,18,18,6] dénote le nombre de neurones dans chacune des couches.

Du côté des modèles RAPDi à base de modèles d'apprentissage profond avec attention (APA)

(architecture présentée à la figure 5.1), 10 architectures ont également été considérées, et celles-ci sont présentées au tableau 6.3. Rappelons que pour cette seconde nature de modèle d'AP, la colonne « Architecture FF » du tableau 6.3 réfère à l'architecture du réseau *feed forward* de la figure 5.1, tandis que N_B réfère au nombre de blocs d'attention utilisés.

Ceci dit, bien que le modèle le plus performant selon la section 5.1.1 (rubans uniformes) était un modèle à base de réseau *feed forward*, rien n'assure que cela est également le cas pour les rubans CFD. Les phénomènes étudiés ici se distinguent par l'influence d'effets 3-D, et donc, la complexité des modèles à base d'apprentissage profond avec attention (APA) peut être requise. Dans le même ordre d'idées, notons que les architectures présentées aux tableaux 6.2 et 6.3 ne sont pas nécessairement les mêmes que celles présentées aux tableaux 5.1 et 5.2 respectivement.

6.4.2 Performances des modèles étudiés

Les performances des modèles listés aux tableaux 6.2 et 6.3 ont été évaluées à partir de la norme de l'erreur euclidienne moyenne en température $\langle e_T \rangle$, en densité de courant $\langle e_{J_s} \rangle$ et en potentiel électrique $\langle e_{V_s} \rangle$, telles que définies à l'équation (5.1). Le tableau 6.4 présente les erreurs commises par les modèles RAPDi à base de réseau *feed forward*, tandis que le tableau 6.5 indique celles associées aux modèles à base d'APA. Rappelons par ailleurs que l'entraînement de modèles d'apprentissage profond est un processus probabiliste.

Tout d'abord, on souligne la performance accrue des modèles de grande taille comparative-ment au chapitre précédent, autant pour les modèles de la famille « NN » que « APA ». Lors du chapitre portant sur les rubans uniformes, le modèle retenu possédait environ 200 000 paramètres. Dans le présent chapitre, un modèle comme « NN-Gros-3 », qui n'est pas le modèle le plus performant, possède déjà plus de 900 000 paramètres. Les modèles de grande taille, souvent mieux adaptés à la modélisation de phénomènes complexes, présentent ici de meilleures performances.

En ce sens, à la lumière des résultats présentés au tableau 6.4, le modèle RAPDi à base de réseau *feed forward* démontrant les meilleures performances est le modèle NN-Gros-4. Celui-ci engendre des erreurs $\langle e_T \rangle$ (13.6 K), $\langle e_{J_s} \rangle$ (1.33 MA cm⁻²) et $\langle e_{V_s} \rangle$ (0.618 V) inférieures à tous les autres modèles de sa catégorie. Par ailleurs, notons que le modèle le plus performant utilise une fenêtre glissante d'une taille de 31 éléments, contrairement aux modèles de rubans uniformes, qui n'utilisaient que 5. Les modèles RAPDi à base de réseau *feed forward* semblent alors favoriser un plus grand « contexte » spatial lors de la modélisation de ruban CFD. Une hypothèse plausible est que la taille de la fenêtre glissante est corrélée à la longueur de transfert (CTL), mais plus d'études doivent être menées afin de confirmer cette affirmation.

TABLEAU 6.4 Erreurs des modèles RAPDi étudiés à base de réseau *feed forward*.

Nom du modèle	$\langle e_T \rangle$ (K)	$\langle e_{J_s} \rangle$ (MA cm ⁻²)	$\langle e_{V_s} \rangle$ (V)
NN-Profond-1	35.2	2.66	1.51
NN-Profond-2	31.6	2.47	1.32
NN-Profond-3	36.7	2.92	1.87
NN-Moyen-1	42.3	4.09	2.49
NN-Moyen-2	28.3	2.27	1.14
NN-Moyen-3	26.4	2.21	1.08
NN-Gros-1	27.4	2.12	1.15
NN-Gros-2	21.5	1.86	0.922
NN-Gros-3	16.1	1.52	0.772
NN-Gros-4	13.6	1.33	0.618
Modèle de référence	178	9.70	11.2

TABLEAU 6.5 Erreurs des modèles RAPDi étudiés à base d'APA.

Nom du modèle	$\langle e_T \rangle$ (K)	$\langle e_{J_s} \rangle$ (MA cm ⁻²)	$\langle e_{V_s} \rangle$ (V)
APA-Profond-1	41.7	2.85	1.63
APA-Profond-2	32.5	2.69	1.40
APA-Profond-3	37.3	3.06	1.73
APA-Moyen-1	29.7	2.34	1.13
APA-Moyen-2	31.0	2.18	0.953
APA-Moyen-3	26.0	1.85	1.02
APA-Gros-1	19.1	1.77	0.767
APA-Gros-2	17.0	1.59	0.659
APA-Gros-3	14.7	1.26	0.455
APA-Gros-4	16.1	1.22	0.524
Modèle de référence	178	9.70	11.2

Du côté des modèles RAPDi à base d'APA, le modèle le plus performant est le APA-Gros-3, selon les résultats obtenus au tableau 6.5, avec des erreurs euclidiennes moyennes $\langle e_T \rangle$, $\langle e_{J_s} \rangle$ et $\langle e_{V_s} \rangle$ de 14.7 K, 1.26 MA cm⁻² et 0.455 V respectivement ainsi qu'une fenêtre glissante de 25 éléments.

Le modèle RAPDi retenu pour la suite de ce mémoire est donc le modèle RAPDi à base de réseau *feed forward* NN-Gros-4. Le modèle NN-Gros-4 possède une erreur $\langle e_T \rangle$ 1.1 K moins élevée que son homologue APA-Gros-3, tandis que le modèle à base d'APA a pour seul avantage d'obtenir des erreurs $\langle e_{J_s} \rangle$ et $\langle e_{V_s} \rangle$ légèrement inférieures au modèle retenu. Notons également que le modèle le plus performant parmi les modèles à base de réseau *feed forward* (NN-Gros-4) est également le modèle de plus grande taille. Ceci suggère alors la possibilité qu'il existe des modèles encore plus performants, puisque le tableau 6.4 n'indique pas la

TABLEAU 6.6 Temps de calcul moyen d’une simulation et facteur d’accélération obtenu par rapport à la MEF 3-D pour différents modèles.

Modèle	Temps de calcul	Facteur d’accélération ($\times$)
MEF 3-D	3.7 h	1
MDF 1-D	0.62 s	21 600
RAPDi 1-D	3.3 s	4000

présence de modèles qui ont surappris. Le surapprentissage est détecté lorsqu’un modèle ayant un plus grand nombre de paramètres que le modèle NN-Gros-4 possède des performances bien inférieures.

En ce qui concerne les performances en temps de calcul, les temps moyens d’une simulation pour la méthode des éléments finis 3-D, le modèle de référence (MDF 1-D) ainsi que le modèle RAPDi (NN-Gros-4) sont présentés au tableau 6.6. La formulation unidimensionnelle des modèles de référence et RAPDi permet d’obtenir les solutions $T(x, t)$ et $J_s(x, t)$ pour un temps de calcul de l’ordre des secondes, tandis que la MEF 3-D nécessite en moyenne 3.7 heures. Plus précisément, la MDF 1-D et le modèle RAPDi obtiennent un facteur d’accélération de 21 600 et 4000, respectivement. Rappelons également que la différence de temps de calcul entre les deux modèles 1-D est expliquée par la présence du modèle d’apprentissage profond au sein du modèle RAPDi.

On remarque également que les facteurs d’accélération présentés au tableau 6.6 sont plus élevés que ceux du tableau 5.5 en raison de la discrétisation utilisée. En effet, le chapitre précédent utilise une discrétisation $(N_x, N_t) = (501, 501)$, tandis que ce chapitre utilise une discrétisation telle que $(N_x, N_t) = (301, 301)$. Les coûts en calcul sauvés lorsqu’on compare au tableau 5.5 sont donc majoritairement expliqués par le nombre réduit d’itérations ainsi que la taille matricielle des systèmes à résoudre (équation (4.18)).

Par ailleurs, le temps d’entraînement du modèle RAPDi est d’environ 1h sur un GPU (GeForce RTX 3060 Ti) pour le jeu données considéré. Bien que les données soient moins abondantes que celles pour les rubans uniformes, l’entraînement d’un modèle RAPDi pour les rubans CFD requiert de plus gros modèles d’apprentissage profond (davantage de paramètres θ) en raison de la complexité des phénomènes étudiés. Le temps d’entraînement pour le jeu de données de rubans uniformes et de rubans CFD est donc similaire ici.

6.5 Solutions électrothermiques d'un ruban CFD

6.5.1 Solution du problème thermique

Le modèle RAPDi possède certes une rapidité de calcul impressionnante, mais l'exactitude des solutions électrothermiques reste à démontrer. La présente section illustre un exemple de solution thermique obtenue à partir du modèle RAPDi (NN-Gros-4). Par ailleurs, comme il s'agit d'un seul exemple, l'annexe C illustre une solution thermique supplémentaire ainsi qu'une solution électrique. La simulation de l'annexe C a spécialement été choisie afin de présenter un cas de figure où le modèle RAPDi reproduit moins fidèlement la MEF 3-D, tout en ayant une performance néanmoins très satisfaisante.

La figure 6.6 illustre les résultats de l'exemple choisi de solution d'un problème thermique d'un ruban CFD le long de la ligne $A - A'$ obtenus par la MEF 3-D, la MDF 1-D ainsi que le modèle RAPDi, à divers instants. D'une part, le modèle RAPDi se superpose très bien avec la solution de la MEF 3-D. On note toutefois une exception pour la figure 6.6a), où une erreur ponctuelle d'environ 1 K apparaît. Par ailleurs, bien que la solution du modèle RAPDi semble présenter une légère déformation à la gauche de ce même endroit ($x \approx 7.5$ cm), celle-ci s'explique simplement par l'absence d'échantillonnage de la MEF 3-D à cet emplacement. Ceci dit, l'imprécision ponctuelle à la figure 6.6a) n'impacte pas la dynamique du profil de température de façon considérable puisque celle-ci disparaît à l'instant suivant ($t = 4$ ms). En ce sens, le modèle RAPDi n'engendre pas une erreur monotone croissante, malgré le fait que l'on étudie un régime dominé par les pertes Joule, comme mentionné précédemment. Du côté des solutions du modèle RAPDi, aux figures 6.6c) et 6.6d), on remarque que la transition entre le régime supraconducteur et le régime normal (T_0 à T_c) ainsi que la vitesse de propagation de la zone normale (NZPV) sont correctement représentées.

Sans surprise, la dynamique thermique du ruban CFD échappe inévitablement au modèle de référence (MDF 1-D). Bien qu'une résistance d'interface effective $R_{inter,J}$ soit utilisée afin de reproduire la CTL du ruban CFD (figure 6.4), le modèle de référence produit une NZPV bien trop lente comparativement à la NZPV réelle. La zone normale n'a alors pas le temps de se propager, résultant en des profils de température $T(x, t)$ trop étroits. Un résultat important est ainsi démontré : la modélisation d'un ruban CFD à partir d'un modèle 1-D nécessite impérativement l'usage de deux résistances d'interface effectives distinctes, soit $R_{inter,J}$ et $R_{inter,T}$. La MDF 1-D souffre donc d'une résistance d'interface effective $R_{inter,T}$ mal calibrée, mais surtout d'une résistance d'interface effective thermique non adaptative en fonction de l'état électrothermique. Dans le même ordre d'idées, la température maximale modélisée par la MDF 1-D aux figures 6.6c) et 6.6d) est supérieure à celle de la MEF 3-D puisque

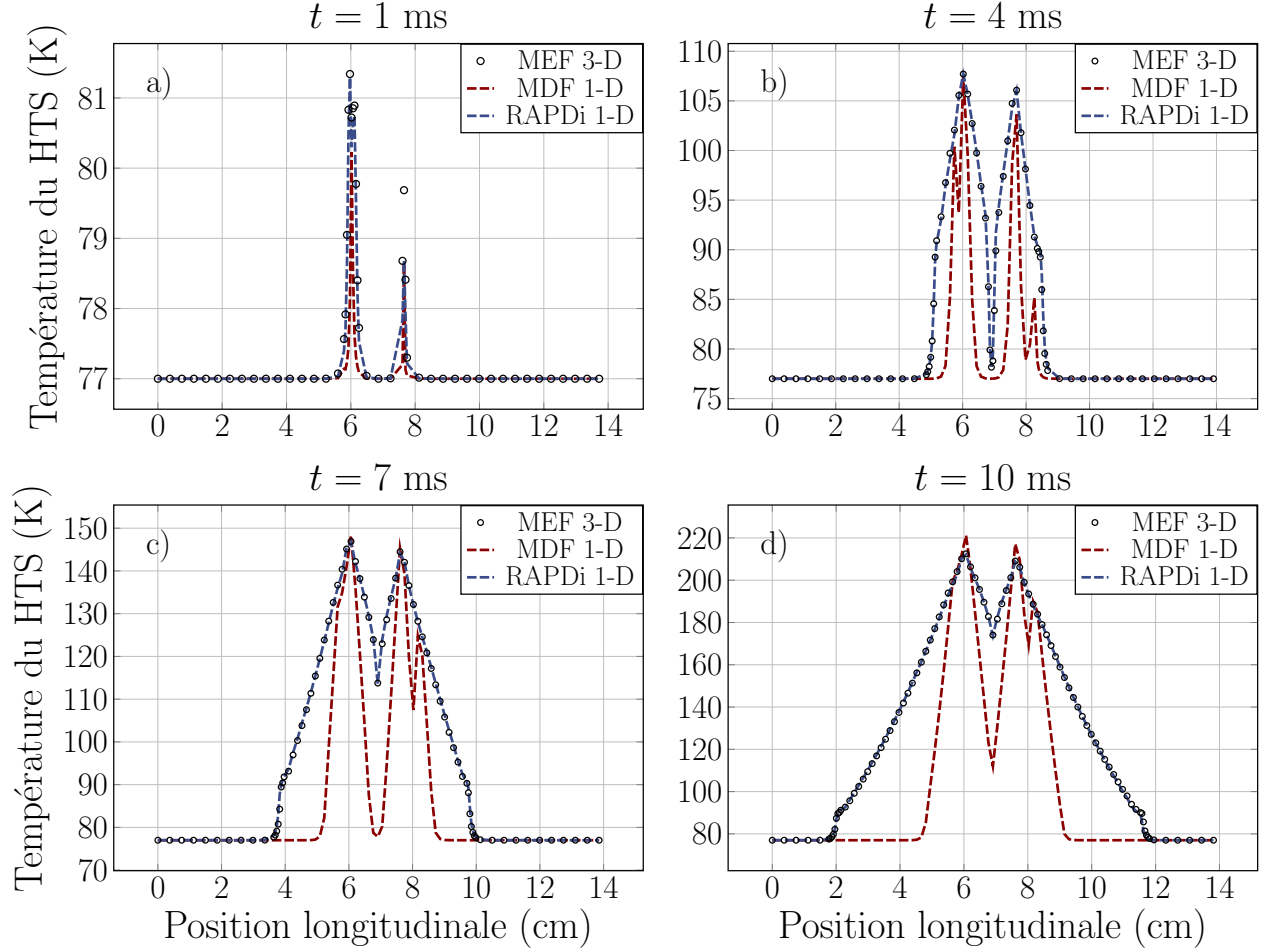


FIGURE 6.6 Température longitudinale à l'intérieur de la couche HTS d'un ruban CFD aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms lors d'un régime dominé par les pertes Joule, simulé par la MEF 3-D, la MDF 1-D (modèle de référence) et le modèle RAPDi. La résistance d'interface effective $R_{inter,J}$ est de $17 \mu\Omega \text{ cm}^2$.

la résistance d'interface effective du problème thermique est plus grande que la résistance d'interface faible du ruban CFD ($R_{inter,T} > R_{inter}$). Par conséquent, la solution thermique associée au modèle de référence produit des erreurs trop significatives pour être une méthode viable, surpassant même les erreurs observées dans le cas d'un ruban uniforme 2-D.

Le modèle RAPDi met donc à profit une résistance d'interface effective $R_{inter,J}$ fixe ainsi qu'une résistance d'interface effective $R_{inter,T}$ adaptative selon le contexte thermique ($T, Q_s, Q_a, Q_{inter}, Q_{conv}$) afin de reproduire la dynamique du quench le long de $A - A'$. À l'opposé, la MDF 1-D est imprécise et incapable de capturer cette même dynamique en raison de la $R_{inter,T}$ fixe qui la constitue. Rappelons que le modèle RAPDi ne fait pas qu'exclusivement modéliser une résistance d'interface effective $R_{inter,T}$ adaptative, il s'agit seulement de l'impact d'un

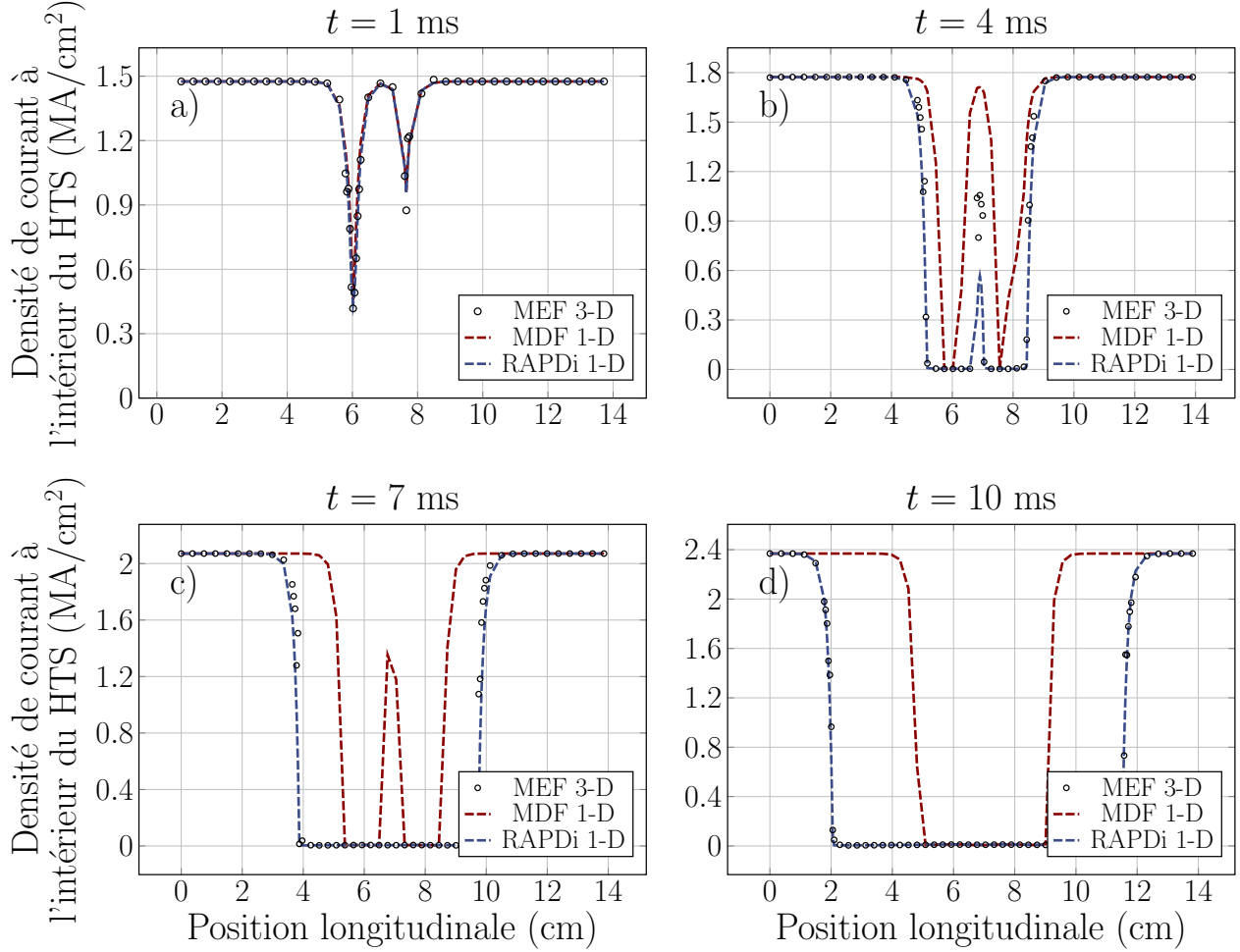


FIGURE 6.7 Densité de courant au milieu de la couche HTS d'un ruban CFD aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms lors d'un régime dominé par les pertes Joule, simulé par la MEF 3-D, la MDF 1-D (modèle de référence) et le modèle RAPDi. La résistance d'interface effective $R_{inter,J}$ est de $17 \mu\Omega\text{cm}^2$.

seul préfacteur (d_θ), tel qu'indiqué à l'équation (6.5).

6.5.2 Solution du problème électrique

La figure 6.7 illustre la solution en densité de courant de la couche HTS $J_s(x, t)$ associée à la solution thermique présentée précédemment le long de la ligne $A - A'$ à quatre instants distincts. Du côté du modèle RAPDi, on observe que ses solutions ainsi que celles de la MEF 3-D coïncident une fois de plus. On remarque toutefois quelques différences aux figures 6.7a) et 6.7b). Plus précisément, une erreur ponctuelle est présente à la figure 6.7a) en raison de l'erreur ponctuelle de 1 K commise par le modèle RAPDi à la figure 6.6a). Quant à la

figure 6.7b), plusieurs points sont omis au centre du ruban ($x \approx 7$ cm) par le modèle RAPDi. Toutefois, on soupçonne que cette erreur s'explique par une tolérance trop grossière lors de la résolution du problème électrique par la MEF 3-D, puisque le modèle RAPDi modélise une solution thermique exacte au même instant (voir figure 6.6b)). L'erreur observée n'est donc pas issue de la résolution du problème thermique. La simulation par éléments finis adopte régulièrement une tolérance relative de 10^{-3} afin de réduire le temps de calcul et aider la convergence numérique. Il est cependant connu au laboratoire qu'une telle tolérance relative n'est pas toujours suffisante, pouvant alors se traduire par des artéfacts numériques dans la solution $J_s(x, t)$ ou une solution localement imprécise. La solution du modèle RAPDi ne souffre pas de ces problèmes puisque le modèle proposé résout l'équation (4.18) avec un critère d'arrêt près de l'erreur machine. Enfin, on observe que le modèle RAPDi modélise fidèlement la longueur de transfert pour les divers instants de la figure 6.7. Cependant, notons que l'emplacement du transfert de courant est légèrement décalé dans l'espace aux instants $t = 4$ ms et $t = 7$ ms, mais ce décalage disparaît à $t = 10$ ms. Ceci dit, il est important de rappeler que le modèle RAPDi 1-D modélise un phénomène électrique 3-D, et ce dernier n'est entraîné qu'à partir de la température T , et non de la densité de courant J_s . Toutefois, modéliser précisément la solution en température $T(x, t)$ permet de reconstruire la solution en densité de courant $J_s(x, t)$ sans accumulation d'erreur importante, tout comme c'était le cas pour les rubans uniformes 2-D.

De son côté, la MDF 1-D souffre de l'inexactitude de sa solution thermique ainsi que d'une accumulation d'erreur importante. Les conséquences de l'imprécision du modèle de référence sont ainsi reflétées au travers des figures 6.7b) à 6.7d), où le partage de courant a lieu au mauvais emplacement pour chacune de ces figures. La MDF 1-D modélise alors une zone normale jusqu'à deux fois moins grande que la taille réelle. C'est notamment ce qui est présenté à la figure 6.7d) : le modèle de référence prédit un transfert de courant décalé de 3 cm par rapport à l'emplacement réel, rendant ainsi cette méthode inapte à modéliser le comportement électrique d'un ruban CFD. En ce sens, ces résultats mettent en évidence une fois de plus la pertinence d'ajouter une résistance d'interface effective adaptative au problème thermique.

Pour ce qui est de la solution en potentiel électrique, la figure 6.8 présente les solutions obtenues le long de la ligne $A - A'$ par la MEF 3-D, la MDF 1-D ainsi que le modèle RAPDi à plusieurs instants. On constate tout d'abord que la solution $V_s(x, t)$ issue du modèle RAPDi surestime la différence de potentiel totale du ruban CFD ($V_s(0, t) - V_s(L, t)$) pour tous les instants de la figure 6.8. Cette surestimation du potentiel électrique est expliquée par le fait que (i) la CTL est souvent légèrement décalée vers l'extérieur par rapport à la MEF 3-D, tel qu'expliqué ci-dessus, et (ii) la solution en densité de courant du modèle RAPDi lorsque

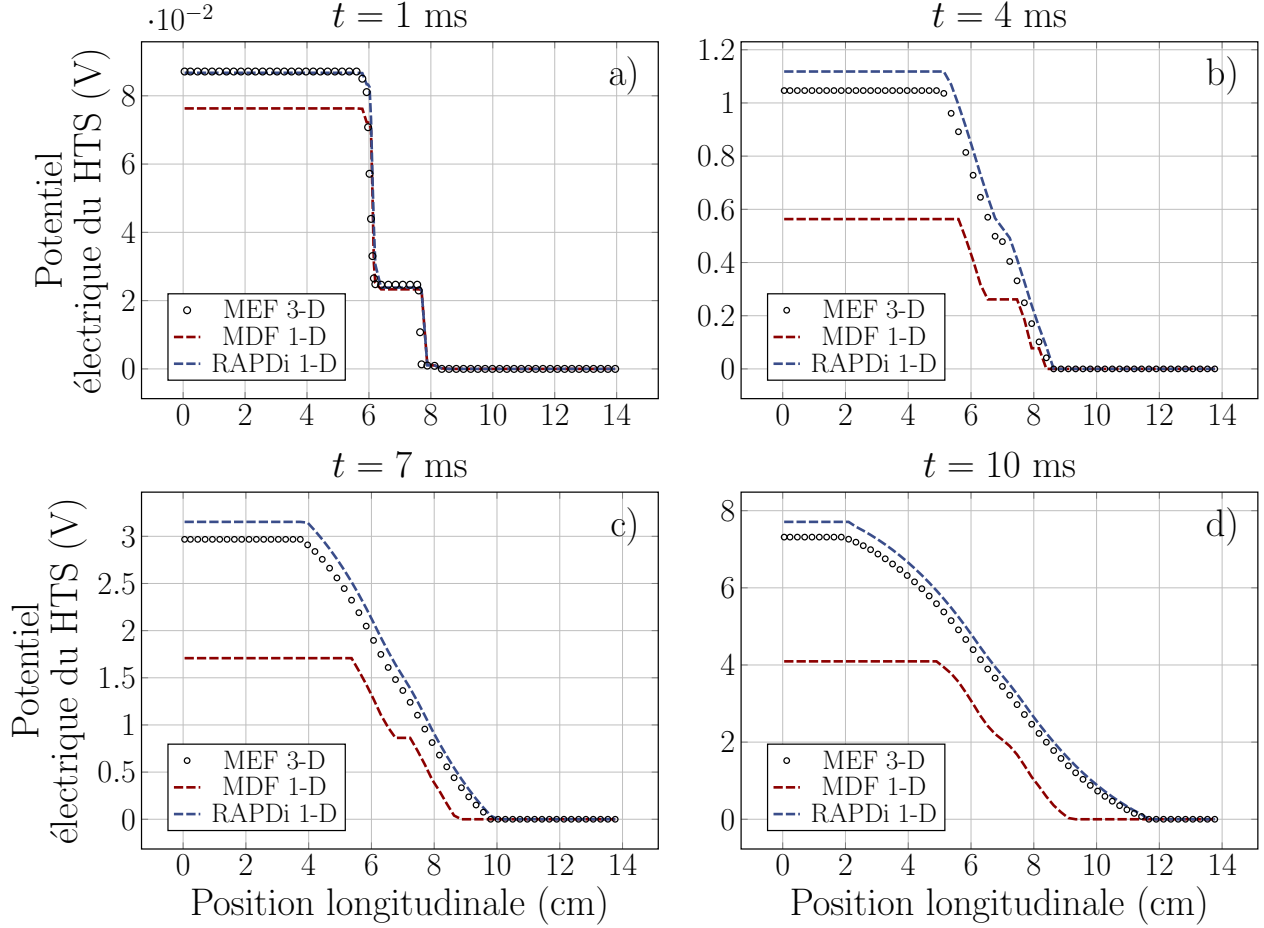


FIGURE 6.8 Potentiel électrique de la couche HTS d'un ruban CFD aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms lors d'un régime dominé par les pertes Joule, simulé par la MEF 3-D, la MDF 1-D (modèle de référence) et le modèle RAPDi. La résistance d'interface effective $R_{inter,J}$ est de $17 \mu\Omega \text{ cm}^2$.

$J_s \approx 0$ a tendance à être légèrement sur-estimée. Le champ électrique à l'intérieur du HTS, exprimé par $E_s = \rho_s J_s$, est par conséquent plus élevé avec le modèle RAPDi qu'avec de la MEF 3-D. Il advient alors que la différence de potentiel électrique l'est également comme le champ électrique est intégré selon l'équation (5.2). Ceci dit, l'erreur relative maximale en potentiel électrique, telle que définie à l'équation (5.3), n'atteint que 5.40 % à $t = 10$ ms, tel que présenté au tableau 6.7. Cette erreur s'élève toutefois jusqu'à 6.84% à $t = 4$ ms pour cette simulation spécifique. L'erreur relative maximale en potentiel électrique peut cependant être inférieure, tel que démontré au tableau C.1 de l'annexe C. Néanmoins, les erreurs commises sur le profil V_s sont majoritairement expliquées par l'hypothèse que le bord du ruban CFD se comporte localement comme un ruban uniforme. Lorsqu'on compare les solutions thermiques du modèle RAPDi à la MEF 3-D, celles-ci présentent une performance sans reproche. Ce

TABLEAU 6.7 Erreur relative maximale de la solution $V_s(x, t)$ pour un ruban CFD à différents instants pour la MDF 1-D et le modèle RAPDi.

Temps	ΔV_{rm} MDF 1-D	ΔV_{rm} RAPDi 1-D
1 ms	12.4%	0.429%
4 ms	46.2%	6.84%
7 ms	42.4%	6.24%
10 ms	44.0%	5.40%

n'est que lorsqu'on résout le problème électrique que les erreurs associées aux hypothèses sont mises en évidence, notamment à travers V_s (figure 6.8), comme celle-ci intègre les quantités ρ_s (équation (2.7)) et J_s (figure 6.7). Ceci dit, le modèle RAPDi présente des performances remarquables pour un modèle 1-D, parvenant à restituer fidèlement la physique locale de ruban CFD 3-D.

Du côté de la MDF 1-D, la solution en potentiel électrique $V_s(x)$ est fortement erronée. L'accumulation d'erreur ainsi que la faible NZPV modélisée rendent les profils de $V_s(x)$ incompatible avec la solution de la MEF 3-D, tels qu'illustré aux figures 6.7b) à 6.7d). La solution $V_s(x, t)$ du modèle RAPDi semble seulement souffrir d'une petite mise à l'échelle, tandis que le modèle de référence produit un profil $V_s(x)$ incohérent en raison de sa distribution de température $T(x, t)$ inexacte. Par conséquent, la solution en potentiel électrique de la MDF 1-D engendre une erreur relative maximale s'élevant à 44.0% à $t = 10$ ms, et atteignant plus de 46.2% à $t = 4$ ms. Ces résultats suggèrent une fois de plus que la méthode des différences finies 1-D est une méthode numérique à éviter lors de la modélisation de ruban CFD.

CHAPITRE 7 CONCLUSION

7.1 Synthèse des travaux

Les travaux de ce mémoire ont permis de développer un tout nouveau modèle numérique à base d'apprentissage profond dont le but est de prédire la dynamique du quench dans les rubans HTS de façon précise et peu coûteuse en temps de calcul. Pour ce faire, le modèle développé repose sur deux stratégies, soit (i) mettre à profit les méthodes d'apprentissage profond développées au cours des dernières années en appui aux méthodes numériques traditionnelles, et (ii) réduire la dimensionnalité du problème, formulant non pas le problème électrothermique en deux ou trois dimensions, mais plutôt en une seule dimension le long du ruban. Une telle réduction de dimensionnalité est possible en raison de la géométrie quasi-linéique propre aux rubans supraconducteurs à haute température critique.

Le modèle « récurrent d'apprentissage profond guidé par la physique basé sur les différences finies » (RAPDi) est donc introduit, combinant le modèle 1-D de Levin *et al.* à un modèle d'apprentissage profond afin de prédire les solutions $T(x, t)$ et $J_s(x, t)$ le long du ruban HTS. Le modèle RAPDi généralise alors l'approche du modèle des coefficients effectifs (développé précédemment au laboratoire) à des problèmes électrothermiques propres à la dynamique du quench, tout en offrant une propriété forte intéressante : l'extension du domaine. Cette propriété, qui s'apparente à celles des réseaux de neurones convolutifs, permet alors d'entraîner le modèle sur une longueur de quelques centimètres, et d'ensuite allonger la longueur du ruban post-entraînement. Cette propriété, qui tient son origine de l'invariance en translation de la méthode des différences finies ainsi que de la fenêtre glissante par rapport au problème électrothermique, permet ainsi de s'affranchir de la limite des modèles éléments finis actuels, soit leur incapacité à modéliser des rubans de grande longueur. Par conséquent, le modèle RAPDi ouvre la porte à des simulations de rubans aux dimensions plus réalistes, élargissant donc le champ d'application des simulations traditionnelles.

Afin de démontrer les performances et les propriétés du modèle proposé, le modèle RAPDi a d'une part été entraîné à partir d'un jeu de données 2-D de rubans uniformes, sur lequel la solution électrothermique a été extraite au centre du ruban. Le modèle RAPDi prédit la solution électrothermique $T(x, t)$, $J_s(x, t)$ ainsi que $V_s(x, t)$ pour une variété de rampes de courant, de régimes thermiques, de résistances d'interface, ainsi que d'inhomogénéités le long de la densité de courant critique. Le modèle proposé prédit alors non seulement la dynamique longitudinale du quench 450 fois plus rapidement que la méthode des éléments finis 2-D, mais la solution obtenue est équivalente à la MEF, contrairement au modèle 1-D classique. Plus

précisément, le modèle RAPDi produit une erreur relative maximale en potentiel électrique de seulement 1.41 %, tandis que celle du modèle de Levin *et al.* est de 26.8%. Par la suite, la propriété d'extension du domaine a été démontrée : un modèle RAPDi entraîné sur des rubans d'une longueur de 7 cm a calculé la solution d'un ruban de 1.4 mètre en seulement 1.6 minute, soit 1110 fois plus efficace que la MEF 2-D. L'erreur relative maximale en potentiel électrique n'est alors que de 1.3%, malgré la longueur importante du ruban. Par ailleurs, des rubans uniformes de plus de 14 mètres (740 inhomogénéités) peuvent être simulés dans ce contexte en seulement 25.5 minutes.

D'autre part, le modèle RAPDi a été entraîné à partir de simulations réalisées sur des géométries de rubans CFD 3-D, soit une architecture de rubans dont le partage de courant est décrit par un phénomène électrothermique tridimensionnel complexe qui échappe aux modèles 1-D classiques. Toutefois, il a été démontré qu'un ruban CFD se comporte localement comme un ruban uniforme le long de ses extrémités. Ce résultat, inconnu jusqu'à présent, permet alors de mettre à profit le modèle développé : le modèle RAPDi modélise la solution le long de l'extrémité du ruban, exploitant ainsi son architecture commune avec le modèle de Levin *et al.* (ruban uniforme). Cependant, certaines modifications sont nécessaires, dont l'introduction de deux nouvelles variables, soit la résistance d'interface effective du problème thermique $R_{inter,T}$ ainsi que la résistance d'interface effective du problème électrique $R_{inter,J}$. La résistance d'interface $R_{inter,T}$ modélise la chaleur générée par la résistance d'interface faible du ruban CFD, tandis que la résistance d'interface $R_{inter,J}$ décrit le comportement électrique le long de l'extrémité du ruban à l'origine de la résistance d'interface R_{CFD} . Une fois ces modifications effectuées et l'architecture optimale du modèle RAPDi est déterminée, les solutions électrothermiques ont été présentées. Du côté du problème thermique, le modèle RAPDi reproduit la transition entre l'état supraconducteur et l'état normal ainsi que la NZPV de façon précise. Le modèle de référence modélise quant à lui une NZPV trop lente, malgré l'ajout de la résistance d'interface effective $R_{inter,J}$, ce qui résulte en des solutions thermiques imprécises et inutilisables. L'utilisation d'une résistance d'interface effective $R_{inter,J}$ pour le problème électrique et d'une résistance d'interface effective $R_{inter,T}$ adaptative pour le problème thermique est donc la clé qui permet modéliser un ruban CFD 3-D à partir d'un modèle 1-D. Du côté du problème électrique, le modèle RAPDi décrit adéquatement la solution en densité de courant, bien que ce dernier n'ait été entraîné que sur des profils de température. Ce n'est que lors de l'étude des solutions en potentiel électrique qu'il est possible d'observer une faible erreur ($\Delta V_{rm} = 5.4\%$) après 10 ms. Ceci dit, le modèle RAPDi reproduit la solution électrothermique du ruban CFD en seulement 3 secondes, tandis que le temps de calcul MEF 3-D s'élève à plus de 3.7 heures, soit une accélération du temps de calcul de plus de 3 ordres de grandeur (4 000).

7.2 Limitations de la solution proposée

Le modèle 1-D proposé dans ce mémoire possède deux caractéristiques limitantes. D’une part, le modèle tient en grande partie ses propriétés du modèle 1-D de Levin *et al.* Par conséquent, le modèle RAPDi est spécialisé pour les géométries quasi-planaires sous la forme de rubans. Ce type de géométrie est requis afin que l’hypothèse d’homogénéisation sous-jacente au modèle soit en mesure d’approximer convenablement le comportement thermique de certaines couches, dont le substrat. Par ailleurs, dans sa forme actuelle, le modèle est uniquement capable de modéliser un partage de courant entre deux couches, soit entre la couche stabilisatrice et la couche supraconductrice. Il est ainsi impossible pour le modèle RAPDi de modéliser des rubans formés d’un empilement de couches stabilisatrices de différentes natures, où le partage de courant a alors lieu entre trois couches ou plus. Enfin, le modèle RAPDi est dépendant du maillage et celui-ci s’applique uniquement aux problèmes électrothermiques, négligeant ainsi la dépendance en champ magnétique B de la densité de courant critique J_c . Cette dépendance peut s’avérer essentielle lors de la nucléation du quench de certaines architectures de rubans.

D’autre part, le modèle RAPDi est limité par le principal inconvénient d’un modèle d’apprentissage profond, soit la sensibilité à l’espace de paramètres défini lors de l’entraînement. En ce sens, il est a priori impossible pour ce dernier d’extrapoler en dehors de cet espace tout en assurant la validité des prédictions effectuées. Si l’on désire utiliser une distribution de $I_c(x)$ réelle, il est donc essentiel de s’assurer que cette distribution respecte l’amplitude des fluctuations de J_c vues durant l’entraînement. Dans le même ordre d’idées, le modèle RAPDi doit définir l’emplacement exact dans l’espace (selon la ligne $A - A'$ ou $B - B'$) où la solution est modélisée avant même l’entraînement. Une fois cet emplacement établi, il est impossible d’obtenir la solution électrothermique ailleurs. Les performances du modèle sont également sensibles à la position de cet emplacement, et donc certains emplacements sont à privilégier, tel que démontré au chapitre 6.

7.3 Améliorations futures

La première étape afin de perfectionner le modèle RAPDi serait de mieux balancer le jeu de données lors de la génération de celui-ci. Dans le cadre de ce mémoire, les données ont toutes été générées selon des lois uniformes continues, où la frontière de l’espace des paramètres a été déterminée arbitrairement. Il est toutefois évident que chacun des paramètres n’a pas le même impact sur la dynamique du quench ou sur la température maximale observée. En ce sens, une piste d’amélioration possible serait de générer les données conjointement avec des

méthodes bayésiennes adaptatives ou d'apprentissage actif afin d'optimiser l'exploitation des données pour un nombre limité de simulations.

Dans le but de répondre à certaines limitations du modèle RAPDi, le modèle proposé peut également être reformulé de façon à modéliser un partage de courant entre plus de deux couches. Pour ce faire, l'équation transcendante du partage de courant doit être dupliquée avec l'ajout d'un terme de couplage en fonction du potentiel électrique de chaque couche. En contrepartie, une telle formulation peut alourdir le temps de calcul du modèle RAPDi en raison de la complexité du nouveau système d'équations. Si l'on désire approfondir encore davantage le modèle RAPDi, une formulation magnétothermique est également possible afin de modéliser des effets inductifs, avec certaines limites inhérentes à un modèle 1-D.

Enfin, une piste de recherche future consisterait à généraliser le problème électrique du modèle RAPDi pour une plus grande gamme de fonctions, c'est-à-dire allant au-delà d'un partage de courant de nature exponentielle. Si un tel objectif est atteint, les erreurs du modèle proposé lors de la modélisation de rubans CFD pourraient être considérablement réduites. Par ailleurs, une seconde piste de recherche, au sens plus large, serait d'appliquer la méthodologie développée dans ce mémoire pour d'autres variétés de géométries. L'application du modèle RAPDi à des géométries cylindriques est notamment intéressante pour la simulation de câble HTS ou de bobines. Des couplages transverses de chaleur entre différentes parties du problème 1-D permettrait par exemple de prendre en considération le couplage thermique entre deux rubans en contact. Ceci dit, ne perdons pas de vue que le modèle RAPDi est certes un modèle électrothermique, mais seul le problème thermique est paramétré. En ce sens, le modèle RAPDi peut être appliqué, en théorie, à différents types de problème, pour autant qu'il existe un problème thermique au sein de la physique modélisée.

RÉFÉRENCES

- [1] C. Yao et Y. Ma, “Superconducting materials : Challenges and opportunities for large-scale applications,” *iScience*, vol. 24, n°. 6, p. 102541, 2021. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S2589004221005095>
- [2] A. Morandi, “State of the art of superconducting fault current limiters and their application to the electric power system,” *Physica C : Superconductivity*, vol. 484, p. 242–247, 2013, proceedings of the 24th International Symposium on Superconductivity (ISS2011). [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0921453412001037>
- [3] P. Tixador *et al.*, “Impact of conductor inhomogeneity on fel transient performance,” *IEEE Transactions on Applied Superconductivity*, vol. 21, n°. 3, p. 1194–1197, 2011.
- [4] F. Sirois et F. Grilli, “Potential and limits of numerical modelling for supporting the development of hts devices,” *Superconductor Science and Technology*, vol. 28, n°. 4, p. 043002, mar 2015. [En ligne]. Disponible : <https://dx.doi.org/10.1088/0953-2048/28/4/043002>
- [5] D. Goodstein et J. Goodstein, “Richard feynman and the history of superconductivity,” *Physics in Perspective*, vol. 2, p. 30–47, 2000.
- [6] H. K. Onnes, “The resistance of pure mercury at helium temperatures,” *Commun. Phys. Lab. University Leiden*, vol. 120, 1911.
- [7] M. Tinkham, *Introduction to Superconductivity : Second Edition*, ser. Dover Books on Physics. Dover Publications, 2004. [En ligne]. Disponible : <https://books.google.ca/books?id=JOQoAwAAQBAJ>
- [8] N. Nibbio, S. Stavrev et B. Dutoit, “Finite element method simulation of ac loss in hts tapes with b-dependent e-j power law,” *IEEE Transactions on Applied Superconductivity*, vol. 11, n°. 1, p. 2631–2634, 2001.
- [9] S. Furtner *et al.*, “Reel-to-reel critical current measurement of coated conductors,” *Superconductor Science and Technology*, vol. 17, n°. 5, p. S281, mar 2004. [En ligne]. Disponible : <https://dx.doi.org/10.1088/0953-2048/17/5/037>
- [10] THEVA Dünnschichttechnik. (2022). [En ligne]. Disponible : <http://www.teva.com>
- [11] C. Lacroix *et al.*, “Normal zone propagation in various rebco tape architectures,” *Superconductor Science and Technology*, vol. 35, n°. 5, p. 055009, mar 2022. [En ligne]. Disponible : <https://dx.doi.org/10.1088/1361-6668/ac56ff>

- [12] J. Gordon *et al.*, “Lattice and electronic specific heat of $\text{YBa}_2\text{Cu}_3\text{O}_7$,” *Physica C : Superconductivity and its Applications*, vol. 162-164, p. 484–485, 1989.
- [13] N. Bagrets *et al.*, “Current transfer length and interfacial resistance between superconductors and metals in commercial REBCO tapes and cables,” *IEEE Transactions on Applied Superconductivity*, vol. 31, n^o. 6, p. 1–8, 2021.
- [14] G. A. Levin, K. A. Novak et P. N. Barnes, “The effects of superconductor–stabilizer interfacial resistance on the quench of a current-carrying coated conductor,” *Superconductor Science and Technology*, vol. 23, n^o. 1, p. 014021, dec 2009. [En ligne]. Disponible : <https://dx.doi.org/10.1088/0953-2048/23/1/014021>
- [15] Z. Stekly et J. Zar, “Stable superconducting coils,” *IEEE Transactions on Nuclear Science*, vol. 12, n^o. 3, p. 367–372, 1965.
- [16] B. Maddock, G. James et W. Norris, “Superconductive composites : heat transfer and steady-state stabilization,” *Cryogenics*, vol. 9, n^o. 4, p. 261–273, 1969.
- [17] F. Gömöry et J. Šouc, “Stability of DC transport in HTS conductor with local critical current reduction,” *Superconductor Science and Technology*, vol. 34, n^o. 2, p. 025005, févr. 2021.
- [18] M. Wilson et R. Wolf, “Calculation of minimum quench energies in rutherford cables,” *IEEE Transactions on Applied Superconductivity*, vol. 7, n^o. 2, p. 950–953, 1997.
- [19] L. Dresner, *Stability of superconductors*. Springer Science & Business Media, 2006.
- [20] C. Lacroix, F. Sirois et J.-H. F. Lupien, “Engineering of second generation HTS coated conductor architecture to enhance the normal zone propagation velocity in various operating conditions,” *Superconductor Science and Technology*, vol. 30, n^o. 6, p. 064004, may 2017. [En ligne]. Disponible : <https://dx.doi.org/10.1088/1361-6668/aa684f>
- [21] C. Lacroix *et al.*, “Normal zone propagation velocity in 2G HTS coated conductor with high interfacial resistance,” *IEEE Transactions on Applied Superconductivity*, vol. 23, n^o. 3, p. 4701605–4701605, 2013.
- [22] X. Wang *et al.*, “Normal zone initiation and propagation in Y-Ba-Cu-O coated conductors with Cu stabilizer,” *IEEE Transactions on Applied Superconductivity*, vol. 15, n^o. 2, p. 2586–2589, 2005.
- [23] F. Trillaud *et al.*, “Normal zone propagation experiments on HTS composite conductors,” *Cryogenics*, vol. 43, n^o. 3, p. 271–279, 2003. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0011227503000444>
- [24] C. Lacroix et F. Sirois, “Concept of a current flow diverter for accelerating the normal zone propagation velocity in 2G HTS coated conductors,” *Superconductor*

- Science and Technology*, vol. 27, n°. 3, p. 035003, jan 2014. [En ligne]. Disponible : <https://dx.doi.org/10.1088/0953-2048/27/3/035003>
- [25] J.-H. Fournier-Lupien, “Étude de la dynamique du quench dans différentes architectures de rubans supraconducteurs à haute température critique,” Thèse de doctorat, École Polytechnique de Montréal, Novembre 2020. [En ligne]. Disponible : https://publications.polymtl.ca/5564/1/2020_Jean-HughesFournier-Lupien.pdf
- [26] F. Sosa-Rey *et al.*, “Adaptive multi-scale electrothermal model of rebco coated conductors embedded in a commercial power system transient simulator,” *Superconductor Science and Technology*, vol. 32, n°. 10, p. 105005, aug 2019. [En ligne]. Disponible : <https://dx.doi.org/10.1088/1361-6668/ab2e67>
- [27] K. Hornik, M. Stinchcombe et H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, n°. 5, p. 359–366, 1989. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/0893608089900208>
- [28] A. Sherstinsky, “Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network,” *Physica D : Nonlinear Phenomena*, vol. 404, p. 132306, 2020. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0167278919305974>
- [29] J. Chung *et al.*, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” 2014. [En ligne]. Disponible : <https://arxiv.org/abs/1412.3555>
- [30] A. Vaswani *et al.*, “Attention is all you need,” 2023. [En ligne]. Disponible : <https://arxiv.org/abs/1706.03762>
- [31] A. Dosovitskiy *et al.*, “An image is worth 16x16 words : Transformers for image recognition at scale,” 2021. [En ligne]. Disponible : <https://arxiv.org/abs/2010.11929>
- [32] S. Ioffe et C. Szegedy, “Batch normalization : Accelerating deep network training by reducing internal covariate shift,” 2015. [En ligne]. Disponible : <https://arxiv.org/abs/1502.03167>
- [33] J. L. Ba, J. R. Kiros et G. E. Hinton, “Layer normalization,” 2016. [En ligne]. Disponible : <https://arxiv.org/abs/1607.06450>
- [34] K. He *et al.*, “Deep residual learning for image recognition,” 2015. [En ligne]. Disponible : <https://arxiv.org/abs/1512.03385>
- [35] N. Srivastava *et al.*, “Dropout : A simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, n°. 56, p. 1929–1958, 2014. [En ligne]. Disponible : <http://jmlr.org/papers/v15/srivastava14a.html>
- [36] A. Karpatne *et al.*, “Theory-guided data science : A new paradigm for scientific discovery from data,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, n°. 10, p.

- 2318–2331, 2017.
- [37] R. Yu et R. Wang, “Learning dynamical systems from data : An introduction to physics-guided deep learning,” *Proceedings of the National Academy of Sciences*, vol. 121, n°. 27, p. e2311808121, 2024. [En ligne]. Disponible : <https://www.pnas.org/doi/abs/10.1073/pnas.2311808121>
 - [38] M. Raissi, P. Perdikaris et G. Karniadakis, “Physics-informed neural networks : A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, p. 686–707, 2019. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0021999118307125>
 - [39] A. G. Baydin *et al.*, “Automatic differentiation in machine learning : a survey,” *Journal of machine learning research*, vol. 18, n°. 153, p. 1–43, 2018.
 - [40] A. S. Krishnapriyan *et al.*, “Characterizing possible failure modes in physics-informed neural networks,” 2021. [En ligne]. Disponible : <https://arxiv.org/abs/2109.01050>
 - [41] L. Lu *et al.*, “Learning nonlinear operators via deeponet based on the universal approximation theorem of operators,” *Nature Machine Intelligence*, vol. 3, n°. 3, p. 218–229, mars 2021. [En ligne]. Disponible : <http://dx.doi.org/10.1038/s42256-021-00302-5>
 - [42] Z. Li *et al.*, “Fourier neural operator for parametric partial differential equations,” 2021. [En ligne]. Disponible : <https://arxiv.org/abs/2010.08895>
 - [43] Z. Hao *et al.*, “Gnot : A general neural operator transformer for operator learning,” 2023. [En ligne]. Disponible : <https://arxiv.org/abs/2302.14376>
 - [44] J.-H. Fournier-Lupien, F. Sirois et C. Lacroix, “A physics-guided recurrent data-driven model for time transient solution of nonlinear partial differential equations subjected to unpredictable external excitations,” 2022, manuscrit non publié. Disponible sur demande à Frédéric Sirois.
 - [45] V. Stanev *et al.*, “Machine learning modeling of superconducting critical temperature,” *npj Computational Materials*, vol. 4, n°. 1, p. 29, June 2018. [En ligne]. Disponible : <https://doi.org/10.1038/s41524-018-0085-8>
 - [46] T. D. Le *et al.*, “Critical temperature prediction for a superconductor : A variational bayesian neural network approach,” *IEEE Transactions on Applied Superconductivity*, vol. 30, n°. 4, p. 1–5, 2020.
 - [47] X. Liu *et al.*, “The feasibility of designing a back propagation neural network to predict the levitation force of high-temperature superconducting magnetic levitation,” *Superconductor Science and Technology*, vol. 35, n°. 4, p. 044004, mar 2022. [En ligne]. Disponible : <https://dx.doi.org/10.1088/1361-6668/ac55f5>

- [48] Z. Ke *et al.*, “Prediction models establishment and comparison for guiding force of high-temperature superconducting maglev based on deep learning algorithms,” *Superconductor Science and Technology*, vol. 35, n°. 2, p. 024005, jan 2022. [En ligne]. Disponible : <https://dx.doi.org/10.1088/1361-6668/ac455d>
- [49] Y. Nakai et S. Noguchi, “Quench prediction for rebco pancake coils using lstm,” *IEEE Transactions on Applied Superconductivity*, vol. 34, n°. 5, p. 1–5, 2024.
- [50] P. Peng *et al.*, “A quench behavior predictive model for high temperature superconducting magnet based on deep-learning neural network,” *IEEE Transactions on Applied Superconductivity*, vol. 35, n°. 5, p. 1–6, 2025.
- [51] A. Sadeghi *et al.*, “Intelligent probability estimation of quenches caused by weak points in high-temperature superconducting tapes,” *Energies*, vol. 16, n°. 1, 2023. [En ligne]. Disponible : <https://www.mdpi.com/1996-1073/16/1/193>
- [52] Y. Touloukian, *Thermophysical properties of matter*. New York : IFI/PLENUM, 1970.
- [53] D. R. Lide, *CRC Handbook of Chemistry and Physics*. New York : CRC Press, 2010.
- [54] S. J. Hagen, Z. Z. Wang et N. P. Ong, “Anisotropy of the thermal conductivity of $\text{YBa}_2\text{Cu}_3\text{O}_{7-y}$,” *Phys. Rev. B*, vol. 40, p. 9389–9392, Nov 1989. [En ligne]. Disponible : <https://link.aps.org/doi/10.1103/PhysRevB.40.9389>
- [55] T. A. Friedmann *et al.*, “Direct measurement of the anisotropy of the resistivity in the a-b plane of twin-free, single-crystal, superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$,” *Phys. Rev. B*, vol. 42, p. 6217–6221, Oct 1990. [En ligne]. Disponible : <https://link.aps.org/doi/10.1103/PhysRevB.42.6217>
- [56] J. Lu, E. S. Choi et H. D. Zhou, “Physical properties of hastelloy® c-276TM at cryogenic temperatures,” *J. Appl. Phys.*, vol. 103, n°. 6, p. 064908, March 2008. [En ligne]. Disponible : <https://doi.org/10.1063/1.2899058>

ANNEXE A DISCRÉTISATION DE L'ÉQUATION DE LA CHALEUR PAR LA MÉTHODE DE DU FORT-FRANKEL

L'équation thermique non linéaire à discrétiser s'exprime ainsi :

$$C(T) \frac{\partial T}{\partial t} + \frac{\partial}{\partial x} \left(-K(T) \frac{\partial T}{\partial x} \right) = Q, \quad (\text{A.1})$$

où $Q = Q_a + Q_s + Q_{inter} + Q_{conv}$, et l'on résout pour la température T . D'une part, on suppose un maillage uniforme défini par $x_i = i\Delta x$ et $t_n = n\Delta t$, avec $T_i^n = T(x_i, t_n)$ et $Q_i^n = Q(x_i, t_n)$. D'autre part, les dérivées première et seconde dans l'espace sont approximées par la méthode des différences finies centrées d'ordre 2, i.e.

$$\frac{\partial T}{\partial x} \approx \frac{T_{i+1}^n - T_{i-1}^n}{2\Delta x} \quad (\text{A.2})$$

et

$$\frac{\partial^2 T}{\partial x^2} \approx \frac{T_{i+1}^n - 2T_i^n + T_{i-1}^n}{\Delta x^2}, \quad (\text{A.3})$$

respectivement. Cependant, le terme T_i^n ne peut pas intervenir explicitement lors de la discrétisation de la dérivée seconde puisque l'on considère le schéma de discrétisation de Du Fort-Frankel. En conséquence, on substitue T_i^n par l'identité (3.9), et l'équation (A.3) devient alors

$$\frac{\partial^2 T}{\partial x^2} \approx \frac{T_{i+1}^n - 2 \left(\frac{T_i^{n+1} + T_i^{n-1}}{2} \right) + T_{i-1}^n}{\Delta x^2} = \frac{T_{i+1}^n - T_i^{n+1} - T_i^{n-1} + T_{i-1}^n}{\Delta x^2}.$$

Notons également que le terme $\frac{\partial}{\partial x}(-K(T) \frac{\partial T}{\partial x})$ de l'équation (A.1) peut se développer tel que

$$\begin{aligned} \frac{\partial}{\partial x} \left(-K(T) \frac{\partial T}{\partial x} \right) &= -\frac{\partial K}{\partial x} \frac{\partial T}{\partial x} - K(T) \frac{\partial^2 T}{\partial x^2} \\ &= -\frac{\partial K}{\partial T} \left(\frac{\partial T}{\partial x} \right)^2 - K(T) \frac{\partial^2 T}{\partial x^2}, \end{aligned} \quad (\text{A.4})$$

où la dérivée en chaîne $\frac{\partial K}{\partial x} = \frac{\partial K}{\partial T} \frac{\partial T}{\partial x}$ a été utilisée. En combinant les résultats précédents, l'équation thermique (A.1) est alors discrétisée selon la méthode de Du Fort-Frankel :

$$C(T_i^n) \left(\frac{T_i^{n+1} - T_i^{n-1}}{2\Delta t} \right) - \frac{\partial K(T_i^n)}{\partial T} \left(\frac{T_{i+1}^n - T_{i-1}^n}{2\Delta x} \right)^2 - K(T_i^n) \left(\frac{T_{i+1}^n - T_i^{n+1} - T_i^{n-1} + T_{i-1}^n}{\Delta x^2} \right) = Q_i^n. \quad (\text{A.5})$$

Il ne suffit alors que de résoudre cette équation algébrique pour T_i^{n+1} . En regroupant chacun des termes en température, on trouve que

$$\begin{aligned} \frac{C(T_i^n)}{2\Delta t} (1 + A_i) T_i^{n+1} = & \frac{C(T_i^n)}{2\Delta t} (1 - A_i) T_i^{n-1} + \frac{K(T_i^n)}{\Delta x^2} (T_{i+1}^n + T_{i-1}^n) \\ & + \frac{\partial K(T_i^n)}{\partial T} \frac{(T_{i+1}^n + T_{i-1}^n)^2}{4\Delta x^2} + Q_i^n, \end{aligned} \quad (\text{A.6})$$

où

$$A_i = \frac{2K(T_i^n)\Delta t}{C(T_i^n)\Delta x^2}. \quad (\text{A.7})$$

Enfin, en divisant chaque côté de l'équation par $\frac{C(T_i^n)}{2\Delta t} (1 + A_i)$, on obtient l'équation thermique discrétisée par Du Fort-Frankel :

$$T_i^{n+1} = \frac{1}{1 + A_i} \left[(1 - A_i) T_i^{n-1} + A_i (T_{i+1}^n + T_{i-1}^n) + B_i (T_{i+1}^n - T_{i-1}^n)^2 + \gamma_i Q_i^n \right], \quad (\text{A.8})$$

où

$$B_i = \frac{\Delta t}{2C(T_i^n)\Delta x^2} \frac{\partial K(T_i^n)}{\partial T} \quad \text{et} \quad \gamma_i = \frac{2\Delta t}{C(T_i^n)}. \quad (\text{A.9})$$

ANNEXE B ALGORITHME DE LA MÉTHODE DE NEWTON–RAPHSON POUR L'ÉQUATION DU PARTAGE DE COURANT

Entrées : Vecteurs x , $\mathbf{T}^n$ et $\mathbf{J}_s^{n-1}$, scalaire I_{tot} , dictionnaire de paramètres P_PDE

Sorties : Vecteurs $\mathbf{J}_s^n$, J_a , ρ_s , ρ_a

- 1: $J_{front} \leftarrow I_{tot}/A_s$
- 2: $\eta \leftarrow -R_{inter} A_s/A_a \cdot h_a$
- 3: $\rho_a \leftarrow \rho_{Ag}(\mathbf{T}^n)$ (via spline)
- 4: $n \leftarrow \begin{cases} n_0, & T_i < T_c, \\ 1, & \text{sinon} \end{cases}$
- 5: $J_c \leftarrow \begin{cases} J_{c0} \frac{T_c-T}{T_c-T_0} \Phi(x), & T < T_c \\ 1, & T \geq T_c \end{cases}$
- 6: $\varepsilon \leftarrow 10, k \leftarrow 0, m \leftarrow N_x - 2$
- 7: $a \leftarrow -\eta/\Delta x^2$
- 8: vecteurs $b, \mathbf{r}^{(k)} \in \mathbb{R}^m$ non initialisés
- 9: **Initialisation :** $J_s \leftarrow \frac{J_{front}}{\mathbf{J}_s^{n-1}[0] + 10^{-8}} \mathbf{J}_s^{n-1}$
- 10: **Tant que** $\varepsilon \geq \text{tol}$ **faire**
- 11: $k \leftarrow k + 1$
- 12: $J_a \leftarrow (I_{tot} - A_s J_s)/A_a$
- 13: $\rho_s \leftarrow \min \left(\max \left(\frac{E_0}{J_c} \left| \frac{J_s}{J_c} \right|^{n-1}, 10^{-16} \right), \rho_{s,n} \right)$
- 14: $\mathcal{S} \leftarrow \rho_a \max(J_a, 0) - \rho_s \max(J_s, 0)$
- 15: $\frac{\partial \mathcal{S}}{\partial J_s} \leftarrow -\frac{A_s}{A_a} \rho_a - n \rho_s$
- 16: $\frac{\partial^2 J_s}{\partial x^2} \Big|_{x=x_i} \leftarrow (J_{s,i-1}^n - 2J_{s,i}^n + J_{s,i+1}^n)/\Delta x^2$
- 17: $b \leftarrow 2\eta/\Delta x^2 + \partial \mathcal{S}/\partial J_s$
- 18: $-\mathbf{r}^{(k)} \leftarrow \eta \frac{\partial^2 J_s}{\partial x^2} - \mathcal{S}$
- 19:

$$\mathcal{J}^{(k)} = \begin{bmatrix} b_2 & a & 0 & \cdots & 0 \\ a & b_3 & a & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & a & b_{N_x-2} & a \\ 0 & \cdots & 0 & a & b_{N_x-1} \end{bmatrix}$$

20: $\Delta J_s \leftarrow \text{thomas}(a, b, -\mathbf{r}^{(k)})$ (Résoudre $\mathcal{J}^{(k)}\Delta J_s = -\mathbf{r}^{(k)}$)
 21: **Mise à jour :** $J_s^{(k+1)}[1 : \text{end}-1] \leftarrow J_s^{(k)}[1 : \text{end}-1] + \Delta J_s$, $\mathbf{J}_s^n \leftarrow J_s^{(k+1)}$
 22: $\varepsilon \leftarrow \|J_s^{k+1} - J_s^k\| / \|J_s^k\|$
Fin
Retourner : $\mathbf{J}_s^n, J_a, \rho_s, \rho_a$

ANNEXE C SOLUTION ÉLECTROTHERMIQUE DE RUBAN CFD OÙ LE MODÈLE RAPDi N'EST PAS EXACT

Le présent annexe présente la solution électrothermique d'une simulation issue de l'ensemble de test pour une géométrie de ruban CFD 3-D. Cette simulation a été spécialement choisie afin de visualiser les performances du modèle RAPDi lorsque celui-ci n'est pas exact. En ce sens, la solution en température, en densité de courant et en potentiel électrique du modèle RAPDi est présentée aux figures C.1, C.2 et C.3, respectivement. Chacune de ces solutions est comparée à la MDF 1-D ainsi que celle de la MEF 3-D. Par ailleurs, les erreurs relatives maximales en potentiel électrique sont présentées à divers instants au tableau C.1.

TABLEAU C.1 Erreur relative maximale de la solution $V_s(x, t)$ pour un ruban CFD à différents instants pour la MDF 1-D et le modèle RAPDi, où ce dernier n'est pas exact.

Temps	ΔV_{rm} MDF 1-D	ΔV_{rm} RAPDi 1-D
1 ms	14.8%	13.9%
4 ms	57.0%	9.19%
7 ms	51.3%	9.14%
10 ms	52.6%	7.86%

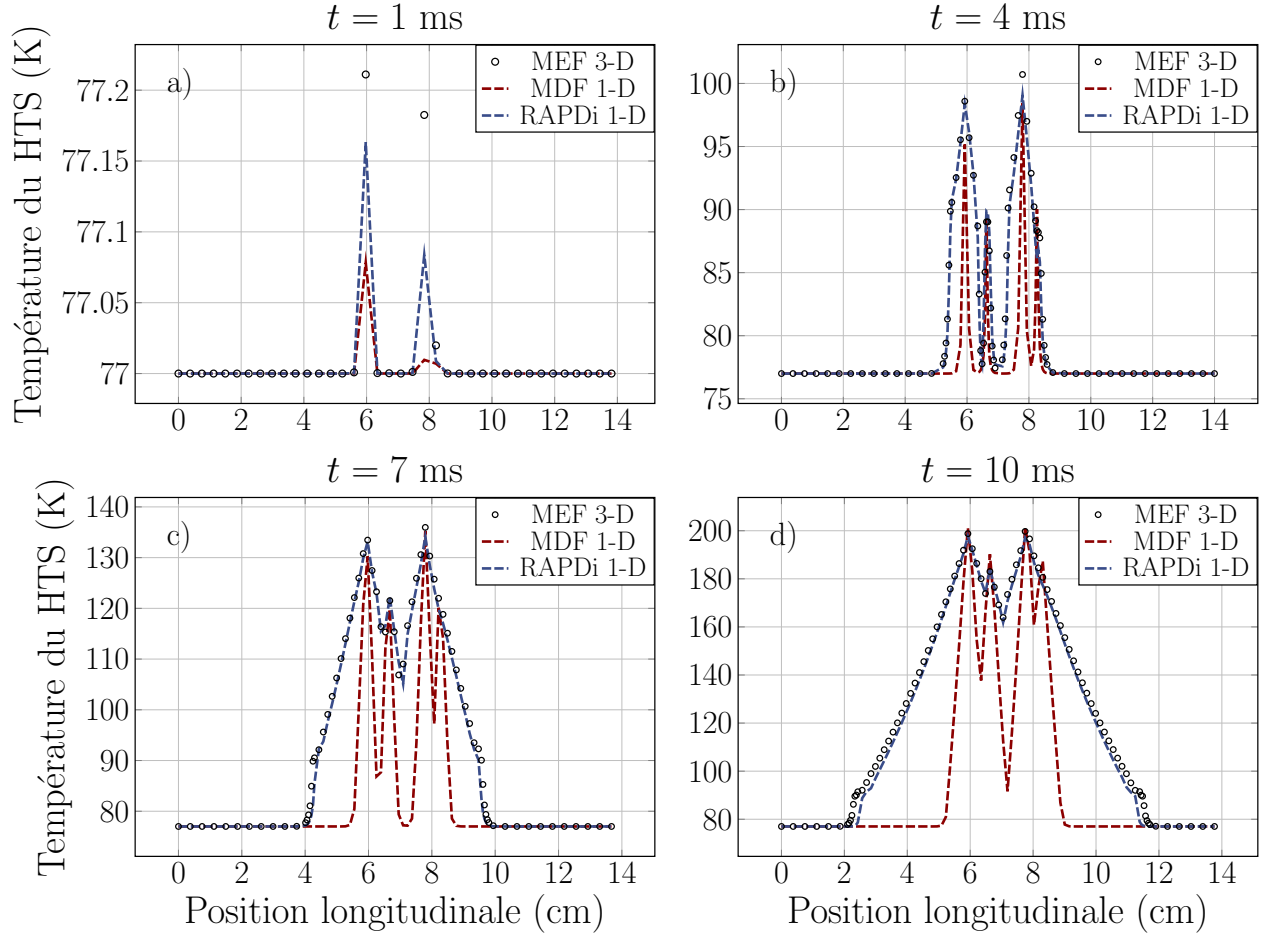


FIGURE C.1 Température longitudinale à l'intérieur de la couche HTS d'un ruban CFD aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms, lors d'un régime dominé par les pertes par effet Joule pour la MEF 3-D, la MDF 1-D (modèle de référence) et le modèle RAPDi, où celui-ci n'est pas exact. La résistance d'interface effective $R_{inter,J}$ est de $8.9 \mu\Omega \text{ cm}^2$.

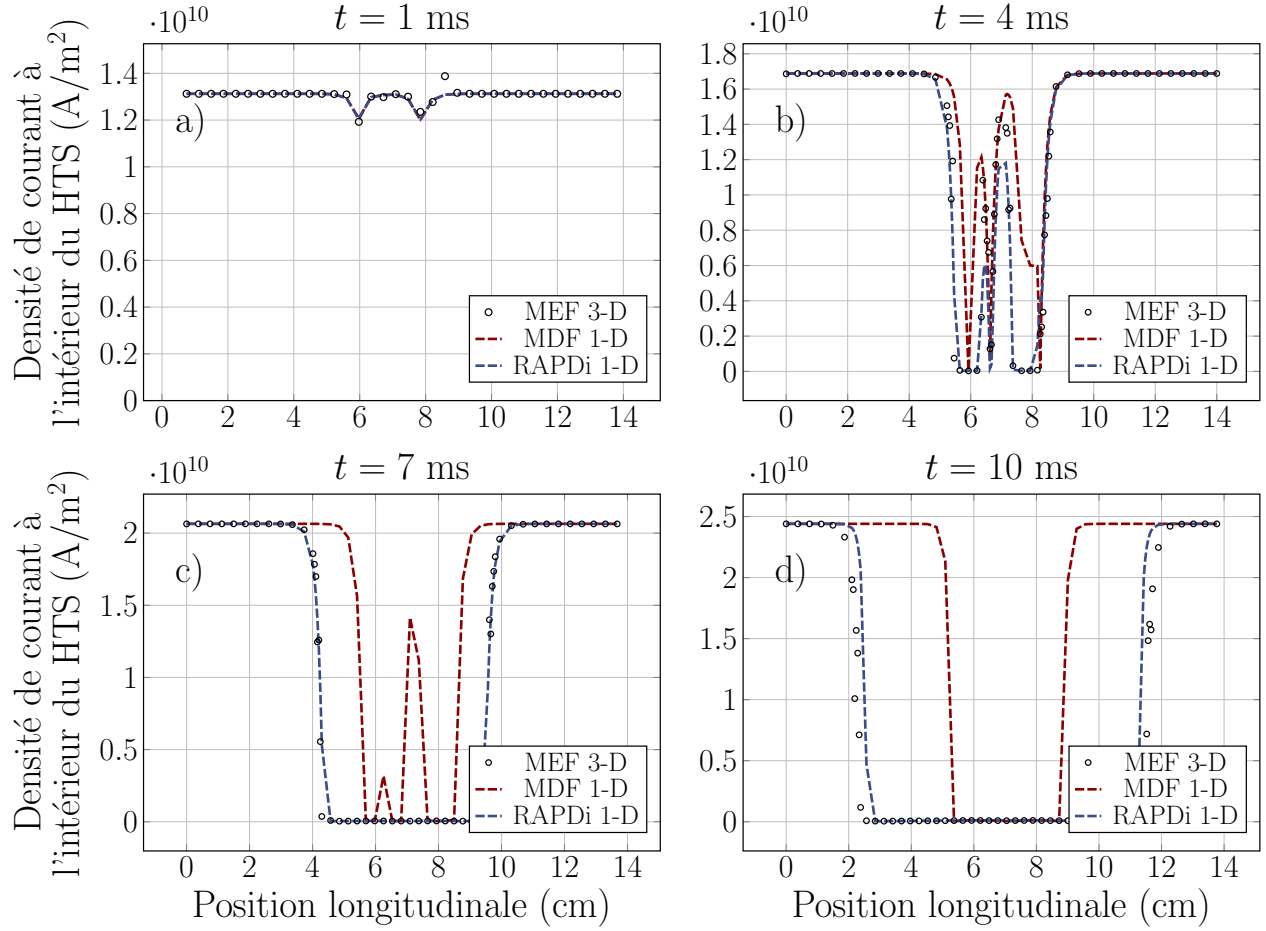


FIGURE C.2 Densité de courant au milieu de la couche HTS d'un ruban CFD aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms, lors d'un régime dominé par les pertes Joule pour la MEF 3-D, la MDF 1-D (modèle de référence) et le modèle RAPDi, où celui-ci n'est pas exact. La résistance d'interface effective $R_{inter,J}$ est de $8.9 \mu\Omega \text{ cm}^2$.

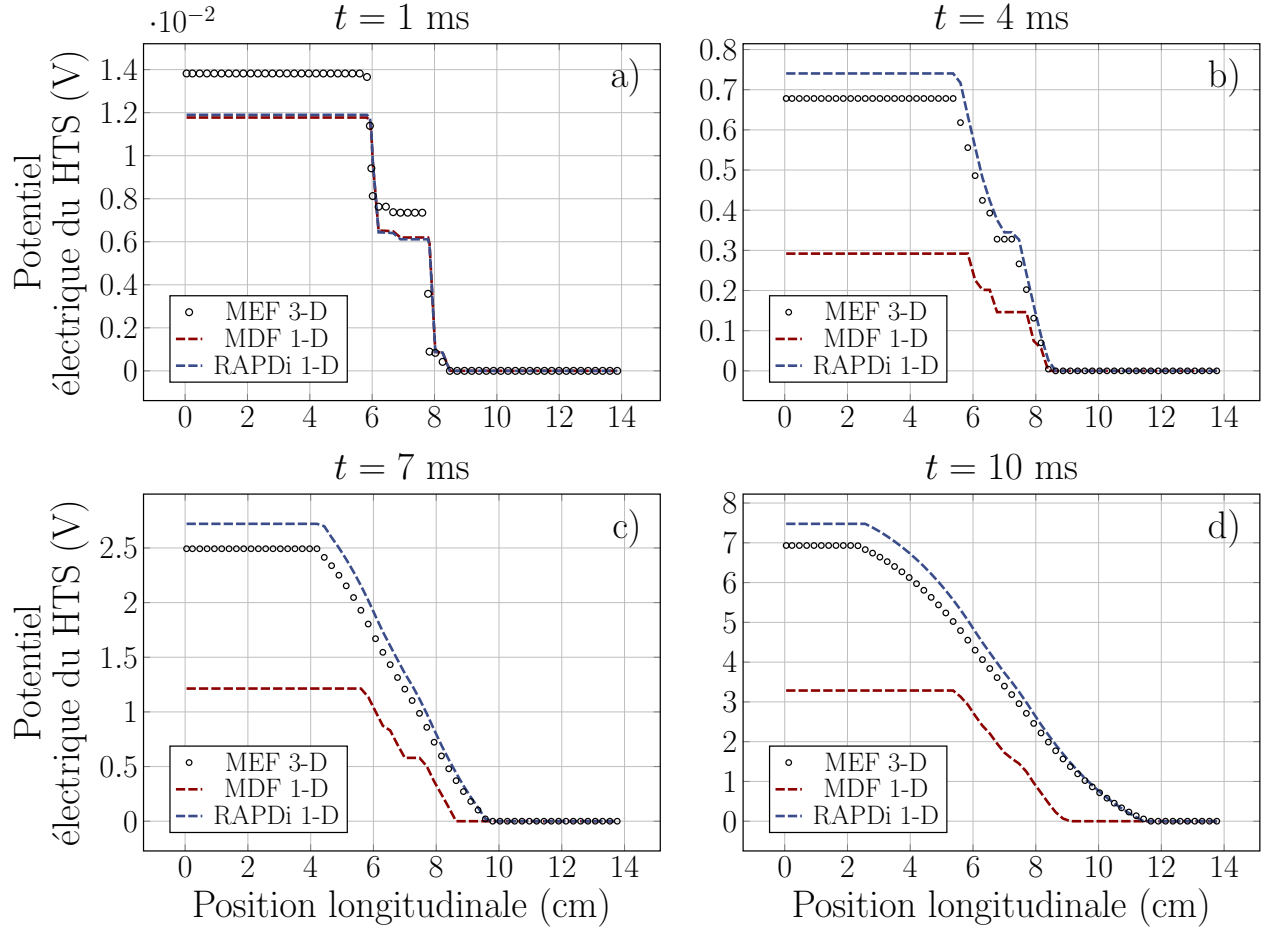


FIGURE C.3 Potentiel électrique de la couche HTS d'un ruban CFD aux instants (a) 1 ms, (b) 4 ms, (c) 7 ms et (d) 10 ms, lors d'un régime dominé par les pertes Joule pour la MEF 3-D, la MDF 1-D (modèle de référence) et le modèle RAPDi, où celui-ci n'est pas exact. La résistance d'interface effective $R_{inter,J}$ est de $8.9 \mu\Omega \text{ cm}^2$.