

Simulation d'événements rares pour des systèmes complexes représentés par des modèles hybrides et Analyse de sensibilité

Contacts EDF : Hassane Chraïbi (hassane.chraibi@edf.fr)

Ce projet de thèse porte sur la fiabilité des systèmes dynamiques complexes représentés par des modèles hybrides i.e. ayant un comportement stochastiques discrets (défaillances, reconfigurations, mécanismes de contrôle et de rétroaction, réparations, etc.) en interaction avec des phénomènes physiques déterministes. L'objectif est de mettre au point, pour ce type de systèmes, des procédures de simulation d'événements rares et d'estimation de leurs probabilités. Ces systèmes peuvent être simulés par l'outil PyCATSHOO développé par EDF. PyCATSHOO adopte le cadre mathématique des processus markoviens déterministes par morceaux (PDMP) introduits par Davis [2] qu'il modélise à l'aide d'automates stochastiques hybrides distribués (DHSA).

Dans une thèse en cours (Thomas Galtier, EDF), deux méthodes de simulation d'événements rares pour les PDMP sont étudiées : l'échantillonnage préférentiel couplé à une méthode de Cross Entropy pour la sélection des paramètres de la densité d'importance [1] et les systèmes de particules en interaction avec une méthode mémorisation [3]. Les PDMPs étant des processus très dégénérés (dégénérés du point de vue stochastique, car ils suivent des trajectoires qui sont déterministes par morceaux), l'implémentation de ces méthodes réclame un soin particulier dans le calcul des rapports de vraisemblance pour la méthode d'échantillonnage préférentiel et dans la procédure d'échantillonnage et de sélection pour la méthode par systèmes de particules en interaction. Il apparaît durant cette thèse deux éléments importants et liés :

1. D'une part, on se rend compte qu'il serait très avantageux d'être capable d'identifier les groupes « minimaux » de composants dont la défaillance conduit à l'événement rare étudié, ce qui n'est pas simple dans un cadre dynamique hybride.
2. D'autre part les méthodes proposées produisent des estimateurs de la probabilité de l'événement rare dont la variance est réduite par rapport à une simulation de Monte Carlo brutale (i.e. plus de précision avec un coût calculatoire moindre). Mais ces estimateurs devraient être enrichis par des études de sensibilité qui permettent de mieux identifier les leviers d'amélioration des performances du système étudié. L'idée est donc de pouvoir produire des informations quantitatives sur les responsabilités, souvent partagées, des différents composants du système dans la réalisation de l'événement rare et ce, également, à un coût calculatoire moindre.

Pour répondre à ces deux questions liées, il serait utile de définir et d'estimer des indices de sensibilité dédiés. Les indices de Sobol sont en général recommandés en analyse de sensibilité [6], mais ils souffrent, étant donnés nos objectifs, de deux défauts :

1. D'une part, ils expriment la sensibilité de la sortie vis-à-vis des entrées du point de vue de la variance, qui est une tendance centrale, et pas du point de vue de l'événement rare, qui est un comportement extrême qui ne concerne que les queues de distribution.
2. D'autre part l'estimation des indices de Sobol peut s'avérer prohibitif du point de vue du temps de calcul, notamment lorsque la dimension (le nombre de composantes ayant un comportement aléatoire) est grande. Pour pallier ces défauts, on peut penser à des méthodes de Morris [7], mais récemment de nouveaux indices (basés sur des moyennes de carrés de dérivées) ont été proposés qui sont liés aux indices de Sobol mais qui sont susceptibles d'être plus robustes vis-à-vis de la dimension [5]. D'autres indices de sensibilité ont été introduits qui eux cherchent à quantifier la sensibilité liée à une quantité d'intérêt, qui pourrait être une probabilité de dépassement de seuil ou un quantile [4].

Le but de cette thèse est d'introduire de nouveaux indices de sensibilité, qui soient adaptés pour les PDMP dans le cadres systèmes complexes, pertinents pour l'objectif de simulation d'événements rares, et pratiques dans le sens où ils puissent être estimés à un coût calculatoire « acceptable ». L'estimation de ces indices pourrait permettre « d'apprendre » les groupes « minimaux » du système hybride et de guider des méthodes d'échantillonnage préférentiel (ou autres) pour obtenir des estimations des probabilités des événements rares présentant des réductions de variance optimales. Si on peut exprimer ces indices comme une espérance d'une fonctionnelle du PDMP, l'estimation de ces indices pourrait elle-même utiliser une méthode de réduction de variance. Ainsi on souhaite in fine mettre au point une méthode séquentielle ou itérative, dans laquelle les indices de sensibilité dédiés et la probabilité d'événement rare sont estimés par une procédure d'échantillonnage préférentiel, de filtrage particulière, ou de Monte Carlo séquentiel.

[1] P. D. Boer, D. Kroese, S. Mannor et R. Rubinstein, A tutorial on the cross-entropy method, *Annals of Operations Research*, 134, 19-67, 2005.

[2] M. H. Davis, *Markov Models and Optimization*, Vol. 49, CRC Press, 1993.

[3] P. Del Moral et J. Garnier, Genealogical particle analysis of rare events, *Ann. Appl. Probab.*, 15, 2496-2534, 2005.

[4] J.-C. Fort, T. Klein et N. Rachdi, New sensitivity analysis subordinated to a contrast, *Communication in Statistics-Theory and Methods*, 45, 4349-4364, 2016.

[5] M. Lamboni, B. Iooss, A.-L. Popelin et F. Gamboa, Derivative-based global sensitivity measures: general links with Sobol' indices and numerical tests, *Mathematics and Computers in Simulation*, 87, 45-54, 2013.

[6] M. D. Morris, Factorial sampling plans for preliminary computational experiments, *Technometrics*, 33, 161-174, 1991.

[7] I. M. Sobol, Sensitivity estimates for nonlinear mathematical models, *MMCE*, 1, 407-414, 1993.