We thank the reviewers for their helpful suggestions to improve the manuscript, " Dynamic Modeling of the Human Coagulation Cascade using Reduced Order Effective Kinetic Models." Below, we have outlined the changes we made to the manuscript to address reviewer suggestions.

Reviewer 2 (Changes highlighted in Yellow on MARKED manuscript)

1) Authors should cite Papadopoulos KP et al. Med. Engr. Physics 36:196 (2014) and compare their 5 ODEs/Logic Rules with the Papadopoulos coarse grained approach of 4 ODEs and 6 parameters.

We thank the reviewer for directing us to the work by Papadopoulos KP et al. This work has been cited.

New Reference:

Papadopoulos, Konstantinos and Gavaises, Manolis and Atkin,Chris (2014) A simplified mathematical model for thrombin generation. Medical Engineering & Physics 36: 196 - 204.

Papadopoulos and co-workers developed a coarse-grained model for thrombin generation that uses 4 Ordinary Differential Equations (ODEs) with 6 parameters. Similar to our work this study uses a reduced order approach in describing a complex biological system with few parameters. However we differ fundamentally in the model formulation as well as in the goals of the study.

Firstly, our model is based on a hybrid approach. In each ODE describing a rate equation we combine traditional enzyme kinetic laws with a logical rule. The enzyme kinetic laws (like Michaelis-Menten) have a physiological basis and the logical rule represents the regulatory activity that controls the rate equation. The product of kinetic laws with logical rules encapsulates a complex interplay between substrates and their regulators that affect the rate equation. Secondly, we have modeled the dynamics of coagulation using the hybrid approach rather than thrombin generation modeled by Papadopoulos KP et al. i.e. our model captures greater mechanistic detail about coagulation and not just thrombin generation. These details allowed predicting the roles of primary inhibitors like ATIII and Protein C. We were also able to use flux analysis to understand how the model balanced initiation, amplification and termination of thrombin generation in both normal and hemophilia modes.

In page 3 lines 76-83 cite this work and compare this model against our hybrid approach

"Recently Papadopoulos and co-workers used a phenomenological mathematical model for thrombin generation [67]. Using a set of four ordinary differential equations they were able to derive an equation for temporal evolution of thrombin generation. However the model focuses more on thrombin generation rather than dynamics of coagulation. Unlike our hybrid approach it does not consider regulatory detail and thus is unable to capture the roles of primary inhibitors like ATIII or protein C. The model also requires that the parameters are adjusted or fine tuned for each data set. Thus in cases where we need to understand the dynamics of coagulation without complex modeling or when there is incomplete mechanistic knowledge, the hybrid

approach is a better candidate. However platelet modeling is a significant advantage of this model over the hybrid approach. Future versions of the hybrid model can involve modeling platelets as done by Papdopoulos et al. in their phenomenological model."

2) How is the particle swarm optimization different from a genetic algorithm? What are the advantages of PSO over a more typical genetic algorithm?

We thank the reviewer for this question. Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) are both population based optimization algorithms. A Genetic Algorithm starts with a randomly generated 'population' that represents the various initial solution vectors of the fitness function (residual in our model). Based on pre-specified fitness selection methods 'parents' are chosen from these population members discarding rest of the population members. The parents are then used to generate 'children' through operations called 'cross-over' and 'mutation'. The fittest children and parents are then used to generate parents for the next generation.

Particle Swarm Optimization like GA starts with a random population of particles that represent the solution vectors. The particles are then allowed to traverse in the search space of the fitness function (residual in our model) with certain velocities. The particle positions (which represent the solutions) are then updated based on a local experience and global experience. These operations are relatively very simple and computationally less expensive as compared to 'cross-over' and 'mutation' that are generally used in GA's. Typical GA is primarily suited for binary value problems rather than continuous real-valued problems. The complexity of crossover and mutation operations in real valued problems increases with increasing precision of the numbers. In addition PSO never discards its solutions like GA. This leads to a greater diversity in solutions and thus allows us to search for a global optimum more effectively especially in case of multi-modal fitness functions.

3) The authors know that APC is only present on endothelium in blood vessels and is not present in blood in tubes in clotting assays, thus PC=0% in Fig. 3. The authors should comment if and how soluble thrombomodulin was used to obtain the data in Fig. 4.

We thank the reviewer for this comment. Physiological concentration of thrombomodulin (Tm) is dependent on endothelial cell surface/volume ratio. Thus the concentration of Tm varies spatially. In a study to understand inhibitory mechanism of protein C on Tissue Factor induced thrombin generation van't Veer and co workers varied the values of Tm from 0-10 nM. In these ranges Tm was found to activate protein C (PC) to activated protein C (APC). The effect of Tm observed in these ranges was found by van't Veer et al. to be potentially physiologically relevant. Based on this study Butenas and co-workers used thrombomodulin concentrations of 0.1 nM and 1 nM to activate protein C. The presence of thrombomodulin at these concentrations did not affect the maximum or total thrombin produced but affected the initiation phase causing a delayed amplification. We replicated the data by Butenas et al. for our study. Our hybrid model as shown in Figures 3,4 & Figure 5 has captured these effects. We added the reference to work by van't Veer and co-workers.

New Reference:

Cornelis van't Veer, Neal J. Golden, Michael Kalafatis and Kenneth G. Mann (1997) Inhibitory Mechanism of the Protein C Pathway on Tissue Factor-induced Thrombin Generation. The Journal of Biological Chemistry 272: 7963-7994.

4) The most common test of thrombin generation conducted in many labs is to add increasing amounts of tissue factor (usually from 0.1 to 10 pM TF). The authors present a Factor VIIa titration (Fig. 7) but it would also be very interesting to present a TF titration since that is the classic test of a models predictive accuracy. Predicting hemophilia is a fairly weak test of any model since severe deficits in FVIII or FIX are so immediately upstream of prothrombinase generation.

We definitely agree with the reviewer's comment that Tissue Factor (TF) is the most common initiator of the coagulation process. In our study we trained our model reproducing the experimental data from Butenas and co-workers. This study used VIIa-TF to initiate the coagulation process in the presence and absence of protein C pathway. VIIa-TF is immediately downstream to TF and is a more potent activator of coagulation than TF alone. Using the same set of parameters from our training data we were able to capture the coagulation dynamics under hemophilia A and hemophilia B conditions initiated using recombinant factor rFVIIa (rFVIIa) from a different group. We believe that this shows a surprisingly remarkable predictive ability and robustness of the model. We would also like to draw the reviewer's attention to the predictions in panel A of Figure 7. We used different levels of trigger 1x-200x to initiate coagulation. Our model currently generalizes the various coagulation triggers or initiators like TF, VIIa-TF or rFVIIa. Thus the model would exhibit the same predictive ability with different levels of TF.

5) The authors could more clearly describe how a specific set of a large number of initial conditions (typically > 10-20 species concentrations in the large ODE models) are processed by the logic rules to give the regulatory term. Since blood kinetics are mostly studied in the tube (no endothelium), Eqn 12 and 13 drop out to result in a 3 ODE model. In this case there are only 3 regulatory terms (v-init, v-amp, and v-inh). Can the author more clearly show how these 3 terms actually depend on various concentrations of actual species that experimentalists might control in the lab.

We kindly acknowledge the reviewer's comment that blood kinetics is studied in the tube and there is absence of endothelium. However Equations 12 and 13 do not drop out in our model in the presence of protein C pathway. These equations are concerned with the inhibitory role of activated protein C. To activate protein C pathway, Butenas and co-workers used soluble thrombomodulin (Tm) at 0.1 nM and 1 nM concentrations. At these concentrations soluble Tm was able to support protein C activation akin to activation of protein C by endothelium bound Tm in vivo.

6) The language should be more uniform in the paper. It seems the terms "regulatory term" and "control term" are used interchangeable.

We thank the reviewer for pointing out this discrepancy. Control term is a function of multiple regulatory factors or functions. We have made the necessary changes to make the language more consistent and avoid any confusion. The new changes are as follows

Line 377 now reads as:

"Each reaction rate was written as the product of two terms, a kinetic term ($\bar{r_j}$) and a control term (ν_j) that depends on multiple regulators: "

Line 380 now reads as:

"Lastly, if an enzyme had both m activating and n inhibitory factors, we modeled the control term as"

7) The quantification of the regulatory role of the platelet is not well described in the paper.

We thank the reviewer for this comment. In our study we do not currently model platelets and cannot comment on their regulatory role in our study. As we have mentioned in our paper platelets are a significant unmodeled component. This is one of the shortcomings of our model. The work by Papadopoulos KP et al. is an excellent way to model platelets in a reduced order model and we intend to model platelets in our model based on this work. However despite this shortcoming we have been able to capture coagulation dynamics in the presence of platelets in hemophilia case as shown in Figure 6.

The model was trained against experimental data from Butenas et al. where they used phospholipid vesicles (PCPS) instead of activated platelets. The ensemble of parameters thus obtained was able to predict the hemophilia case in presence of platelets by a different experimental group. Thus the regulatory role of platelets seems to have been implicitly captured by the kinetic and regulatory terms.

Reviewer 3 (Changes highlighted in Green on MARKED manuscript)

1. Abbreviation TF should be expanded in line 59?.

   We thank the reviewer for pointing out this discrepancy. In line 59, TF has been expanded to Tissue Factor. The line now reads as

   'Initially, thrombin is produced upon cleavage of prothrombin by fluid phase activated factor X (FXa), which itself has been activated by Tissue Factor/factor VII (TF/FVIIa) [10]'

2. Fig. 2. In a figure body function f_j(Z, k) should depend on \kappa according to Methods but does not on k. This should be reconciled.

   We thank the reviewer for the comment. We have made the necessary changes in the methods section in Equation 9 and lines 382 and 383 to show that $f_j(Z, k)$ depends on k but not on kappa.

3. Authors should explain an advantage of their logic approach vs. simple Machaelis-Menten equation taking account several inhibitors (competitive, non-competitive, allosteric) and activators.

   We thank the reviewer for raising a very pertinent point. The hybrid approach uses a combination of logical rules represented by transfer functions along with traditional enzyme kinetic expressions. The use of logical rules is quite advantageous especially when there are several regulators and/or when the exact mechanisms behind regulatory activity are unknown.

   Firstly, using transfer functions simplifies the regulatory logic when several inhibitors and activators are involved. For example, in our study v_{init} is a regulatory term that affects the rate of initiation. This term is influenced by the presence of coagulation activators like Tissue Factor (TF), factor VII (VIIa), factor XI, and factor XII and inhibitors like Tissue Factor Plasminogen Inhibitor (TFPI). Using Michaelis-Menten (MM) equation for each of these activators and inhibitors greatly increases the number of parameters (with each MM equation requiring at the least 3 parameters) that are involved in the regulatory logic. This considerably increases the computational burden in parameter estimation and model simulation.

   Secondly, unlike coagulation where the biochemistry is well defined, in many biological systems there is scarce information about the exact mechanism of enzyme activity (competitive, non-competitive, allosteric etc.). This causes an impediment in determining the exact form MM equation. Transfer functions do not necessarily require a detailed or exact knowledge of the enzyme mechanism. Consider again the example of v_{init} in our study. In Eq 16 we model this as a function of trigger and TFPI. Trigger here can be TF, TF/VIIa, rFVIIa or any other coagulation initiators. Further we use the min/max rule as shown in Equations 6-8 to quantify the cumulative effect of all the triggers and inhibitors (TFPI) on

initiation. Using a logical rule allowed us to reduce the effect of different initiators to a simple rule. This can be especially a very useful and effective method when there is incomplete information. We used Hill like functions to describe the effect of initiators and inhibitors. Each function requires just 2 parameters unlike an MM equation, which requires a minimum of 3 parameters. Transfer functions also have the flexibility of form. As mentioned by us in our study (Page 20, Lines 331-334) they can be a Hill like function or otherwise. Thus the use of logical rules using transfer functions reduces the computational complexity and also helps in encoding unknown/less known regulatory logic.

4. Detailed information on experimental assays should be added in caption of Figs. 3,4, and 5 or in the text.

We thank the reviewer for this comment. Detailed information on experimental assays has been added in the text. Please see page 7, lines 146-159 and the figure legends for Figures 3,4 and 5.

Page 7, lines 146-159 now read as

"In these experiments thrombin generation was initiated by FVIIa-TF using mean plasma concentrations of coagulation proteins and inhibitors. To prepare FVIIa-TF, TF (0.5 nmol/L) was relipidated into 400 μmol/L of phospholipid vesicles (PCPS) by incubation in 20 mmol/L HEPES, 150 mmol/L NaCl, and 2 mmol/L $CaCl_2$ pH 7.4 (HBS/Ca^{2+}) for 30 minutes at 37 °C. The relipidated TF was incubated with 10 pmol/L factor VIIa for 20 minutes to allow the formation of FVIIa-TF. Factors V, VIII and thrombomodulin (Tm) (when protein C activation is required) were added to FVIIa-TF complex. Thrombin generation was then initiated by adding equal volumes of this mixture with a mixture containing prothrombin, factor IX and factor X, TFPI, AT-III and protein C (added when required), protein S (added when required) and factor XI (added when required). In the experimental training data sets that we used for parameter estimation 5 pmol/L FVIIa-TF was used along with 200μmol/L of phospholipid vesicles (PCPS) to initiate thrombin generation. When protein C pathway was involved, protein C and protein S were at mean plasma concentrations and 0.1 nmol/L Tm was used. All the other coagulation proteins and inhibitors i.e. factors X, IX, V, and VIII, prothrombin, TFPI and AT-III were at their mean plasma concentration levels."

5. The authors should show how they introduced time-delay parameter T_D into the model (10)-(14) and clarify whether they solved ODEs or differential-delay equations.

We thank the reviewer for this comment. The time delay parameter T_D is an experiment specific parameter that we introduced to account for the variability in initiation time across training data sets while using a single ensemble of parameters. We solved the reduced order model with 5 differential equations and thereafter introduced the time delay parameter to account for the initiation time in each experiment. Thus T_D remained constant for all data sets in a given experiment but was allowed to vary from experiment to experiment.

6. Table of the final kinetic parameters obtained as a result of fitting procedure should be added to the text to allow readers to reproduce the results.

> We thank the reviewer for this comment. All the parameter ensembles for every experimental simulation can be located at
>
> https://github.com/jeffreyvarner/HybridCoagulationModel_v1
>
> For every experimental simulation the ensemble of parameters are saved in Ensemble.dat. We have provided detailed instructions in README.md on how to run the simulations using the model code.

7. Table with the names of species x_i and their initial concentration should be added.

> We thank the reviewer for this comment. The names of the species are mentioned in line 351 where we state x=(fII,FIIa,PC,APC,ATIII)^{T}. The names of all the species and their initial concentrations for every simulation can be found at:
>
> https://github.com/jeffreyvarner/HybridCoagulationModel_v1/blob/master/Simulations.py
>
> The Simulations.py file has initial concentration for every species in each experiement.

8. Line 344: Commonly the maximum rate is defined as the product of catalytic rate, k, and enzyme concentration, e, or this is enzyme activity instead of author's notation k_{max} and activity \epsilon.

> We thank the reviewer for this comment. We have defined k_j^{max} as the maximum rate for a reaction j and $\in_i$ as the scaled enzyme activity, which catalyzes reaction j. Hence reaction rate $\bar{r}_j$ is the product of maximum rate and scaled enzyme activity. This notation differs from the usual notation where we express maximum rate as the product of catalytic rate k_{cat}*$\in_0$ where $\in_0$ is enzyme concentration. Thus k_j^{max} here is equal to k_{cat}*$\in_0$.

9. Eq. 9: there should be f_i(Z) rather than f_i(x)

> We thank the reviewer for this comment. $f_i(x)$ has been changed to $f_i(Z_i)$ in equation 9. Please see the corresponding equation.

10. Discussion seems to be reduced. Instead of the discussion of reduced model expansion, it would be better to compare briefly the results obtained in the author's reduced approach with the results obtained in the unreduced models reviewed by authors in Introduction.

The reviewers point is well taken. Discussion has been increased to compare the results from the reduced approach with those from mechanistic models. Please refer to Page 19, lines 324-332 for expanded discussion.

"The performance of the proof of principle coagulation model was impressive given its limited size. The most detailed mechanistic model of coagulation is the one by Luan and co-workers built with 193 proteins and protein complexes that are interconnected by 301 reactions [24]. Akin to the hybrid model this mechanistic used the normal thrombin data and hemophilia data from Allen *et al.* [66]. However unlike the mechanistic model we used this data for validation rather than training. Results from our model are surprisingly comparable to the training simulations of the mechanistic model. The initiation time and amplification of thrombin signal were accurately predicted in the normal case whereas in the hemophilia case we correctly predict the initiation time but slightly over predict the amplification signal. This performance is nearly similar to that of the detailed mechanistic model with 301 reactions."

Typos:

Page 6, line 2 there should be "Michaelis-Menten "

The typo has been corrected. Please see the change on Page 6, line 2.