A Review of Physically Based and Thermodynamically Based Constitutive Models for Soft Materials

In this paper, we review constitutive models for soft materials. We specifically focus on physically based models accounting for hyperelasticity, visco-hyperelasticity, and damage phenomena. For completeness, we include the thermodynamically based visco-hyperelastic and damage models as well as the so-called mixed models. The models are put in the frame of statistical mechanics and thermodynamics. Based on the available experimental data, we provide a quantitative comparison of the hyperelastic models. This information can be used as guidance in the selection of suitable constitutive models. Next, we consider visco-hyperelasticity in the frame of the thermodynamic theory and molecular chain dynamics. We provide a concise summary of the viscohyperelastic models including specific strain energy density function, the evolution laws of internal variables, and applicable conditions. Finally, we review the models accounting for damage phenomenon in soft materials. Various proposed damage criteria are summarized and discussed in connection with the physical interpretations that can be drawn from physically based damage models. The discussed mechanisms include the breakage of polymer chains, debonding between polymer chains and fillers, disentanglement, and so on. [DOI: 10.1115/1.4047776]

Keywords: constitutive modeling of soft materials, mechanical properties of soft materials
1 Introduction

Soft actuators [1–5], robotics [6–11], sensors [12,13], and flexible electronics [14,15] are rapidly developing fields that require accurate constitutive models to predict the mechanical behavior of soft materials. A variety of constitutive models have been proposed to model soft material behaviors. These models can be categorized as phenomenological, physically based, and mixed models. The physically based models are referred to as those derived from the structure and deformation mechanisms at the microscopic length scale. Thus, the macroscopic material constants are directly related to the material physical parameters at the microscopic scale. This is an attractive feature of the physically based models as compared to their phenomenological counterparts. Figure 1 illustrates schematically some physically relevant mechanisms at the microscopic length scale. The deformation of molecular chains and the associated entanglements can be linked to the hyperelasticity of soft materials [16–20]. The viscosity of soft materials is known to be related to the free chains [21–25], while their damage is attributed to the breaking of chains and cross-linkers with applied deformation [26–30].

There are several excellent reviews of constitutive models that aim at different mechanical behaviors. Boyce and Arruda [31] and Marckmann and Verron [32] have reviewed phenomenological and physically based hyperelastic models. Drapaca et al. [33] reviewed the nonlinear viscoelastic models based on the fading memory assumption. Wineman [34] reviewed linear and nonlinear viscoelastic phenomenological models for elastomer and soft biological tissues. Banks and Kenz [35] reviewed viscoelastic models based on spring-dash models and Boltzmann superposition laws. Diani et al. [36] summarized the comprehensive experimental observations of Mullins effect caused by damage of elastomeric materials, and the corresponding phenomenological and physically based models were reviewed.

In this paper, we aim to provide a detailed summary and illustrations for the constitutive modeling of soft materials with a specific focus on the physically based models. In particular, we systematically review constitutive models that can capture the essential mechanical behaviors of soft materials such as hyperelasticity, visco-hyperelasticity, and damage. While we focus on the physically based models, for completeness, we also discuss the mixed models, and thermodynamically based viscohyperelastic and damage models.

The remainder of the paper is organized as follows. In Sec. 2, we review the existing hyperelastic models in the context of the deformation mechanisms at the microscopic scales. The viscohyperelastic models and the underlying thermodynamics are discussed in Sec. 3. The damage models are summarized in Sec. 4. The paper is concluded with some remarks and discussion.

The terminology used in this paper follows the work [37], and it is summarized in Table 1 below for the convenience of the reader.

2 Hyperelasticity

2.1 Theoretical Background for Hyperelasticity. In this section, we review the basic theory of thermodynamics and statistics for hyperelastic models.

2.1.1 Thermodynamics for Molecular Chains. In the context of hyperelastic theory [37], the Clausius–Duhem inequality is given by

$$S = \frac{1}{2} C - W = 0$$

where $S$ is the second Piola–Kirchhoff stress tensor (PK2 stress), $C$ is the right Cauchy–Green deformation tensor (the related deformation tensors are briefly recalled in the Appendix), $W$ is the elastic strain energy density function (SEDF). Consequently, PK2 stress can be determined hyper-elastically by

$$S = 2 \frac{\varepsilon W}{\varepsilon C} - p C^{-1}$$

where $p$ is an unknown Lagrange multiplier due to the incompressibility constraint (note that we assume that the soft materials are incompressible). The choice of a suitable SEDF $W$—which fully defines the behavior of the soft material—is the key to the accurate modeling of soft materials. As described in the sequel, such SEDFs can be derived by the consideration of statistical mechanics for molecular chains.

2.1.2 Statistical Mechanics for Molecular Chains. In soft materials, the molecular chains are constrained by their neighboring chains. This confinement decreases the allowed configurational numbers of molecular chains, thus increasing their strain energy. The constraint effects can be modeled by the tube model [38], as shown in Fig. 1(d). The ends of the molecular chain in the tube are fixed at $R$ and $R'$. The lateral motion of the molecular chain is confined in the tube, and the situation is identical to the movement of the chain under an external potential $V[r(s)]$, which is infinite outside of the tube and zero inside of the tube, i.e.,

$$V[r(s)] = \begin{cases} 0 & \text{inside the tube} \\ \infty & \text{outside the tube} \end{cases}$$

where $r(s)$ are the actual configurations of a chain. In this case, the Hamiltonian function can be expressed as

$$H = \frac{3}{2b} k_B T \int_0^l ds \frac{\varepsilon^2 r(s)}{\varepsilon s} + V[r(s)]$$

where $k_B$ is the Boltzmann constant, $T$ is the Kelvin temperature, and $b$ is the length of Kuhn monomer (Fig. 2).

---

2Soft materials are typically modeled as incompressible; this assumption significantly simplifies the analysis, however, it also implies that the media cannot support longitudinal elastic waves.
The partition function for the molecular chain is given by
\[ Z = \int_{\mathbb{R}^3} \delta(\mathbf{r}(s)) e^{-\mathcal{H}/k_B T} \]

\[ = \int_{\mathbb{R}^3} \delta(\mathbf{r}(s)) \exp \left[ -\frac{1}{2k_B T} \int_0^1 \left( \frac{\partial^2 \mathbf{r}(s)}{\partial s^2} \right)^2 + V(\mathbf{r}(s)) \right] \]  

(5)

The partition function \( Z \) has the same expression form as the Green function \( G(\mathbf{R}, \mathbf{R}', N) \), which is the number of Kuhn monomers of a single chain as shown in Fig. 2, it satisfies the diffusion equation [39]

\[ \left[ \frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial R^2} + V(R) \right] G(\mathbf{R}, \mathbf{R}', N) = \delta(R - R') \delta(N) \]  

(6)

with the boundary condition
\[ G(\mathbf{R}, \mathbf{R}', N) = 0 \]  

(7)

Equation (6) has been solved by Edwards and Freed [40] as

\[ G(\mathbf{R}, \mathbf{R}', N) = \prod_{\alpha=x,y} g_\alpha(R_\alpha, R'_\alpha; N) \]  

(8)

where

\[ g_\alpha(R_\alpha, R'_\alpha; N) = \frac{1}{2} \sin \left( \frac{\pi R_\alpha}{d} \right) \sin \left( \frac{\pi R'_\alpha}{d} \right) \exp \left( -\frac{\pi^2 N b^2}{6d^2} \right) \]  

(9)

\[ g_z = \left( \frac{3}{2\pi N b^2} \right)^{1/2} \exp \left( -\frac{3(R_z - R'_z)^2}{2N b^2} \right) \]  

(10)

Table 1 A summary of the terminology

| \( W \) | Elastic strain energy density function (SEDF) |
| \( C \) | The right Cauchy–Green deformation tensor |
| \( S \) | The second Piola–Kirchhoff stress tensor (PK2 stress) |
| \( \Sigma \) | Cauchy stress |
| \( B \) | The left Cauchy–Green deformation tensor |
| \( k_B \) | The Boltzmann constant |
| \( T \) | The Kelvin temperature |
| \( R \) | The length of Kuhn monomer |
| \( n \) | The number density of molecular chains |
| \( N \) | The number of Kuhn monomers of a single chain |
| \( I_1 \) | The first invariant of the right Cauchy–Green deformation tensor |
| \( \lambda_i \) | The principal stretches of deformation gradient tensor |
| \( \Psi_L \) | Langevin distribution function |
| \( L \) | Langevin function |
| \( R \) | Inverse Langevin function |
| \( \lambda_{\text{max}} \) | Maximum stretch |

Fig. 1 The schematic diagram of the molecular chain network with applied deformation: (a) the molecular chain network in the undeformed state, (b) the molecular chain network in the deformed state, (c) the molecular chain network after removing external loading, and (d) the tube model capturing the constraint effects of a single chain caused by its neighboring chains.

Fig. 2 A single chain formed by \( N \) Kuhn monomers with the length \( b \).
in which, \( d \) is the effective diameter of the tube.

The boundary conditions (the schematics depicted in Fig. 1(d)) are

\[
R = \left( \frac{d}{2}, \frac{d}{2}, 0 \right), \quad R' = \left( \frac{d}{2}, \frac{d}{2}, L_0 \right)
\]

where \( L_0 \) is the length of the tube. Substituting this equation into Eqs. (8)–(10), we obtain

\[
Z = G(R, R'; N) = \left( \frac{3}{2\pi Nb^2} \right)^{1/2} \exp\left( -\frac{3L_0^2}{2Nb^2} \right) \exp\left( -\frac{\pi^2 N b^2}{3d^2} \right)
\]

where \( \alpha_0 \) is a parameter that depends on the shape of the cross-section of the tube. The SEDF is formulated by Cho [41] as

\[
W = -k_BT \ln \left( \frac{3}{2\pi Nb^2} \right)^{1/2} \exp\left( -\frac{3L_0^2}{2Nb^2} \right) \exp\left( -\alpha_0 \frac{Nb^2}{d^2} \right)
\]

Thus, we have

\[
W = -k_BT \ln \left( \frac{3}{2\pi Nb^2} \right)^{1/2} \exp\left( -\frac{3L_0^2}{2Nb^2} \right)
- k_BT \ln \left( \frac{2}{d} \right)^2 \exp\left( -\alpha_0 \frac{Nb^2}{d^2} \right)
\]

where the first and the second terms of the SEDF correspond to the deformation of molecular chains in the cross-linked network (Fig. 3(a)) and entanglement constraints (EC) induced by the neighboring chains (Fig. 3(b)), respectively. Considering the limiting chain extensibility, the first term of the Eq. (15) can be replaced by the Langevin probability function as

\[
W = -k_BT \ln \left( C \exp\left( -\frac{R}{b} \beta \left( \frac{R}{Nb} \right) - Nb \ln \frac{\beta R}{Nb} \right) \right)
- k_BT \ln \left( \frac{2}{d} \right)^2 \exp\left( -\alpha_0 \frac{Nb^2}{d^2} \right)
\]

where \( R \) is the magnitude of the end-to-end vector \( R \) (Fig. 2). To obtain specific strain energy from Eq. (16), one needs to connect the macroscopic deformation to the deformation of end-to-end vector \( R \) and the effective diameter of the tube \( d \). This connection is a key issue in formulating a physically based hyperelastic constitutive model. From Eq. (16), we know that the SEDF can be decoupled into two parts in an additive form as

\[
W = W_c + W_e
\]

where \( W_c \) and \( W_e \) correspond to the deformation of molecular chains in the cross-linked network and entanglement constraints induced by the neighboring chains, respectively. Equation (17) gives us the basis to decompose manually the hyperelastic network into two parts: cross-linked network and entanglement network, as shown in Fig. 3.

2.2 Physically Based Models. Physically based models are derived based on the microscopic deformation of molecular chains in the network of soft materials. From Eq. (16), we know that these models differ from each other depending on how the evolution of end-to-end vector \( R \) and the effective diameter of the tube \( d \) are connected to the macroscopic deformation. We classify physically based models into two types:

1. Models neglecting the entanglement constraints, i.e., \( W = W_c \).
2. Models incorporating the entanglement constraints, i.e., \( W = W_c + W_e \).

Fig. 3 The hyperelastic network for soft materials consists of (a) cross-linked network and (b) entanglement network [19] (Reprinted with permission from Elsevier © 2018)
2.2.1 Models Without Entanglement Constraints. Neo-Hookean model: Neo-Hookean model is of the most widely used models with a compact and simple form of SEDF. Based on the Gaussian distribution function, Treloar [42] derived the neo-Hookean model in the following form:

\[ W = \frac{1}{2} G (I_1 - 3) \]  

with a single parameter \( G = n k_B T \) referred to as the initial modulus, where \( n \) is the density of molecular chains, \( I_1 \) is the first invariant of the right Cauchy–Green deformation tensor. Neo-Hookean model provides fairly accurate results in the range of moderate deformation level under tensile, simple shear and biaxial test conditions (not exceeding 50% deformation [32]).

Three-chain model: Based on the assumption of even distribution of the molecular chains along with three principal directions, James and Guth [43] proposed the SEDF by using the Langevin distribution function

\[ W = -G \left( \frac{1}{3} \sum_{i=1}^{3} \ln \left[ \Psi_L \left( b \sqrt{N \lambda_i} \right) \right] \right) \]  

where \( \lambda_i (i = 1,2,3) \) are the principal stretches of deformation gradient tensor, and \( \Psi_L \) is the Langevin distribution function with specific form as

\[ \Psi_L \left( b \sqrt{N \lambda} \right) = C_L \exp \left[ \frac{R}{b} \left( -N \ln \frac{\beta}{\sinh \beta} \right) \right], \]

\[ \beta = L^{-1} \left( \frac{\lambda}{\sqrt{N}} \right) \]  

\[ L(x) = \coth (x) - 1/x \]

where \( C_L \) is a normalized parameter. Arruda–Boyce model: Arruda and Boyce [44] proposed a widely used model as

\[ W = -G \ln \left[ \Psi_L \left( b \sqrt{N \Lambda} \right) \right] \]  

\[ \Lambda = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2} \]  

Here, three-chain and Arruda–Boyce models have two parameters \( G \) and \( N \) that should be determined by experiment.

2.2.2 Models With Entanglement Constraints. Slip-link model: Ball et al. [45] proposed a model taking into account the entanglement constraints in the following form

\[ W = \frac{1}{2} G_e \frac{1}{3} \sum_{i=1}^{3} \lambda_i^2 + \frac{1}{2} G_c \sum_{i=1}^{3} \frac{(1 + \eta \lambda_i^2)}{1 + \eta \lambda_i^2} + \ln \left[ 1 + \eta \lambda_i^2 \right] \]  

where \( G_e = n k_B T \) and \( G_c = n k_B T \) are the moduli of the cross-linked network and entanglement network, respectively, \( n_e \) is the density of molecular chains in the cross-linked network, \( n_e \) is the density of molecular chains in the entanglement network, and \( \eta \) is a material parameter. Based on this work, Edwards and Vilgis [16] considered the limiting chain extensibility and proposed the following SEDF

\[ W = \frac{1}{2} G_e \left[ \frac{1}{3} \left( 1 - \alpha^2 \right) \sum_{i=1}^{3} \lambda_i^2 - \frac{1}{3} \right] + \ln \left[ 1 - \alpha^2 \sum_{i=1}^{3} \lambda_i^2 \right] \]  

\[ + \frac{1}{2} G_c \sum_{i=1}^{3} \left[ \frac{\lambda_i^2 (1 + \eta) (1 - \alpha^2)}{1 + \eta \lambda_i^2} + \ln \left[ 1 + \eta \lambda_i^2 \right] \right] \]  

\[ + \frac{1}{2} \left( 1 - \alpha^2 \right) \sum_{i=1}^{3} \lambda_i^2 \]  

where a new parameter \( \alpha \) is introduced to characterize the inhomogeneous deformation and the limiting chain extensibility.

Extended tube model: Utilizing the tube model [46], Kaliske and Heinrich [17] proposed a physically based model which considers the entanglement constraints as

\[ W = \frac{1}{2} G_e \left[ \frac{1}{3} \left( 1 - \delta^2 \right) \sum_{i=1}^{3} \lambda_i^2 - 3 \right] + \ln \left[ 1 - \delta^2 \left( \sum_{i=1}^{3} \lambda_i^2 - 3 \right) \right] \]  

\[ + \frac{2 G_c}{r^3} \sum_{i=1}^{3} \left( \lambda_i^2 - 1 \right) \]  

where \( G_e \) and \( G_c \) are the parameters with physical meaning as mentioned earlier. \( \delta \) and \( r \) are introduced to characterize finite extensibility and release of topological constraints respectively.

ABGI model: Meissner and Matějka [18] proposed a physically based model by combining the Arruda–Boyce model and Extended tube model. This model can characterize the limiting chain extensibility and the entanglement constraints simultaneously. The stress–stretch relationship is given as

\[ \sigma^{\text{ABGI}}_i = G^{\text{ABGI}}_e \frac{1}{3} \frac{1}{N} \sum_{i=1}^{3} \lambda_i^2 - 9N - \frac{2 G^{\text{ABGI}}_c \lambda_i^2 \rho}{\gamma} + p \]  

where \( G^{\text{ABGI}}_e, G^{\text{ABGI}}_c, \) and \( N \) are the parameters which could be related to microscopic quantities, \( \gamma \) is the parameter with physical meaning, and \( p \) is an unknown scalar resulting from the incompressibility, which can be obtained from the macroscopic boundary conditions.

Micro-sphere model: Miehe et al. [25] proposed a model by assuming that the molecular chains are distributed evenly on the surfaces of a sphere and considered the inhomogeneous deformation of molecular chains. The SEDF is

\[ W = -G^{\text{MS}}_e \ln \left[ \Psi_m \left( \frac{1}{4 \pi} \int \frac{1}{|X|} \left( \frac{X}{|X|} \right)^p \, d|X| \right)^{1/p} \right] \]  

\[ + G^{\text{MS}}_c N u \frac{1}{4 \pi} \int \psi(X) d|X| \]  

where \( G^{\text{MS}}, u, p, q, \) and \( N \) are material constants. The symbol \( u \) stands for the tube geometry parameter, while \( p \) and \( q \) do not have any physical meanings. \( X_0 \) and \( X \) are the end-to-end vectors in the initial and current configurations, respectively.

Non-affine network model: Davidson and Goulbourne [20] developed a model to describe the constitutive behavior of the rubber-like materials under large deformation. This model also considers the
limiting chain extensibility and entanglement constraints simultaneously. They derived the SEDF as

\[ W = \frac{1}{6} G_c I_1 - G_e \lambda_\text{max}^2 \ln (3 \lambda_\text{max}^2 - I_1) + G_e \sum_{i=1}^{3} \left( 1 - \frac{\lambda_i - 1}{\lambda_i} \right) \]  

where \( G_c \) and \( G_e \) are the moduli of the cross-linked network and entanglement network, respectively; the symbol \( \lambda_{\text{max}} \) denotes the maximum stretch of soft materials.

Xiang et al. model: By accounting the limiting chain extensibility and the entanglement constraints, Xiang et al. [19] developed a general hyperelastic model by utilizing the tube model [46], incorporating the deformation of the tube diameter in the model. The SEDF is given as

\[ W = G_c N \ln \left( \frac{3N + \frac{1}{2} I_1}{3N - I_1} \right) + G_e \sum_i \frac{1}{\lambda_i} \]  

where \( G_c \) and \( G_e \) are the moduli of the cross-linked network and entanglement network, respectively.

2.3 Summary 1. We summarize the considered hyperelastic models in Table 2.

<table>
<thead>
<tr>
<th>Without EC</th>
<th>With EC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Treloar, 1943); (James and Guth, 1943); (Arruda and Boyce, 1993)</td>
<td>(Ball et al., 1981); (Edwards and Vilgis, 1986); (Kaliske and Heinrich, 1999); (Meissner and Majl̨ka, 2003); (Miehe et al., 2004); (Davidson and Goulbourne, 2013); (Xiang et al., 2018)</td>
</tr>
</tbody>
</table>

The experimental data used for the fitting of the models (Table 2) are based on (i) uniaxial tension, (ii) pure shear, and (iii) equibiaxial deformation tests reported by Treloar [47]. It is preferable to be able to obtain all the material constants from a single experimental setting, for example, from relatively simple uniaxial tensile tests. Unfortunately, the obtained material constants (from the single experimental setting) may not produce accurate predictions for other experimental settings (pure shear, or equibiaxial deformation) of the identical materials. To illustrate these different capabilities of the constitutive models, two different data sets are used to obtain the material parameters of these constitutive models.

1. The parameters are obtained only by fitting the uniaxial data of Treloar [47]. Then, we plot the stress-stretch curves for pure shear and equibiaxial deformation based on the uniaxial data fitting constants. The corresponding stress-stretch curves are shown in figures labeled with (a) from Figs. 4–12 [47].

2. The parameters are obtained by using the combination of the uniaxial, pure shear and equibiaxial data of Treloar [47] simultaneously. The corresponding stress-stretch curves are shown in figures labeled by (b) in Figs. 4–12.

We use the least square method to extract the material parameters. We depict the predictions of these models with the data of Treloar [47], as shown in Figs. 4–12. The results indicate that the models without the entanglement constraints cannot accurately predict pure shear, and equibiaxial results based only on uniaxial data. We note that the Arruda–Boyce model shows the best performance among the non-EC-models (models without the entanglement constraints).

The fitting results also show that the EC-models predict the behavior more accurately as compared with the non-EC-models if the parameters are obtained only by fitting the uniaxial data; the only exception is the micro-sphere model.

2.3.2 Comparison Among These Constitutive Models. To provide a quantitative comparison between these models, the coefficient \( R^2 \) is introduced and calculated for each model as

\[ R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2} \]  

where \( y_i \) are a set experiment data, which are associated with the predicted value of models \( f_i \). The quantity \( \bar{y} \) represents the mean of the experimental data. Here, the material parameters extracted from uniaxial data and only the pure shear and equibiaxial data are used to calculate the coefficient of determination \( (R^2) \).

Figure 13 summarizes the performance of the models in terms of the coefficient \( (R^2) \). The coefficients \( (R^2) \) for different models are summarized in Table 3. (The other four widely used models [48–52] are also concluded in Table 3 for completeness.)
Fig. 5 Comparison of the engineering stress–stretch between the three-chain model and the data of Treloar [47]: (a) material parameters are extracted by uniaxial data and (b) by uniaxial, pure shear, and equibiaxial data simultaneously.

Fig. 6 Comparison of the engineering stress–stretch between the Arruda–Boyce model and the data of Treloar [47]: (a) material parameters are extracted by uniaxial data and (b) by uniaxial, pure shear, and equibiaxial data simultaneously.

Fig. 7 Comparison of the engineering stress–stretch between the slip-link model and the data of Treloar [47]: (a) material parameters are extracted by uniaxial data and (b) by uniaxial, pure shear, and equibiaxial data simultaneously.
Fig. 8  Comparison of the engineering stress–stretch between the extended tube model and the data of Treloar [47]: (a) material parameters are extracted by uniaxial data and (b) by uniaxial, pure shear and equibiaxial data simultaneously

Fig. 9  Comparison of the engineering stress–stretch between the ABGI model and the data of Treloar [47]: (a) material parameters are extracted by uniaxial data and (b) by uniaxial, pure shear, and equibiaxial data simultaneously

Fig. 10  Comparison of the engineering stress–stretch between the micro-sphere model and the data of Treloar [47]: (a) material parameters are extracted by uniaxial data and (b) by uniaxial, pure shear, and equibiaxial data simultaneously
3 Visco-Hyperelasticity

3.1 The Basic Theory for Visco-Hyperelasticity. In this section, we review the basic theory of thermodynamics for viscohyperelastic theory [53,54].

3.1.1 General Thermodynamic Theory. Any thermodynamic process should satisfy the Clausius–Duhem inequality [53]

\[ S \frac{1}{2} \dot{\mathbf{C}} - \dot{W} \geq 0 \] (32)

The SEDF can be denoted as

\[ W = W(C, \xi_1, \ldots, \xi_n) \] (33)

where \( \xi_a \) (\( a = 1, \ldots, n \)) are the internal variables, substituting Eq. (33) into Eq. (32), we have

\[ \left( S - 2 \frac{\partial W}{\partial \dot{C}} \right) \frac{1}{2} \dot{\mathbf{C}} - \sum_{a=1}^{n} \frac{\partial W}{\partial \dot{\xi}_a} \dot{\xi}_a \geq 0 \] (34)

From Eq. (34), we have

\[ S = 2 \frac{\partial W}{\partial \dot{C}} \] (35)

\[ - \sum_{a=1}^{n} \frac{\partial W}{\partial \dot{\xi}_a} \dot{\xi}_a \geq 0 \] (36)

In order to determine the internal variables, \( n \) set of internal evolution equations should be given as

\[ \dot{\xi}_a = \xi_a(C, \xi_1, \xi_2, \ldots, \xi_n) \] (37)

Equations (35)–(37) are fundamental equations for the dissipation processes. The viscoelastic process is also known as a dissipation process, so it should satisfy these equations. Therefore, the key problem for formulating a thermodynamically based model is to choose a reasonable SEDF \( W \), internal variables \( \xi_a \), and their evolution equations. There are no general expressions for

Fig. 11 Comparison of the engineering stress–stretch between the non-affine network model and the data of Treloar [47]: (a) material parameters are extracted by uniaxial data and (b) by uniaxial, pure shear and equibiaxial data simultaneously

Fig. 12 Comparison of the engineering stress–stretch between the Xiang et al. model and the data of Treloar [47]: (a) material parameters are extracted by uniaxial data and (b) by uniaxial, pure shear and equibiaxial data simultaneously
viscohyperelastic models. Any constitutive equation would be regarded as reasonable once it satisfies Eq. (36). There are two ways to construct the evolution equation: the first way is to construct it based on experimental observations, while an alternating way is led by the structure of molecular chains. The basic theories acknowledging the structure of molecular chains will be discussed in Sec. 3.1.3 in detail.

3.1.2 Thermodynamic Theory for the Standard Linear Solid Model. The standard linear solid model is a widely used rheological model. There are two kinds of representations for the standard linear solid model: Maxwell representation and Kelvin representation, as shown in Figs. 14 and 15. Here, we only review the basic theory of the Maxwell representation since it is used by many researchers [53,55–62].

For the Maxwell representation, the deformation of branches $A$ and $B$ is equal to the applied macroscopic deformation, i.e., $\mathbf{F} = \mathbf{F}_A = \mathbf{F}_B$. For branch $B$, the deformation gradient tensor can be multiplicatively decomposed into two parts as $\mathbf{F} = \mathbf{F}_E\mathbf{F}_v$ (more details about multiplicative decomposition can be found in Refs. [53,55,63]), where $\mathbf{F}_E$ and $\mathbf{F}_v$ are the deformation gradient tensors of the spring and dashpot in branch $B$, respectively.
as shown in Fig. 14. So, the SEDF $W$ in Eq. (33) can be formulated as
\[ W = W_1(C) + W_2(C_e) \]  
(38)
where $C = F^T F$ and $C_e = F_e^T F_e$ are the right Cauchy–Green deformation tensor of springs $A$ and $B$ (Fig. 14), respectively. Substituting Eq. (38) into Eq. (32), we obtain
\[ (S - S^E - F^{-1}S F^{-T}) : \frac{1}{2} C + C_s S_e : D_e \geq 0 \]  
(39)
where $D_e = \text{sym}(L_e)$, $F_s$ and $L_e$ are the deformation gradient tensor and the velocity gradient tensor of viscous damper, respectively (shown in Fig. 14). The PK2 stress of the spring $A$ ($S^E$) and the spring $B$ ($S_e$) are defined as
\[ S^E = 2 \frac{\partial W_1(C)}{\partial C} \]  
(40)
and
\[ S_e = 2 \frac{\partial W_2(C_e)}{\partial C_e} \]  
(41)
From Eq. (39), we have
\[ S = S^E + F^{-1}S F^{-T} \]
\[ C_s S_e : D_e \geq 0 \]  
(42)
Various specific constitutive equations could be obtained by selecting different forms of the SEDF and the evolution laws of $D_e$. The derived constitutive equation would be reasonable if it satisfies Eq. (42) and agrees with the specific experimental observations.

### 3.1.3 Basic Theory of Molecular Chain Dynamics.

In this subsection, we give a brief introduction to two widely used basic theories of molecular chain dynamics: Rouse and reptation model [38,41]. It should be pointed out that further readings on this subject can be found from the monograph of Cho [41].

The Cauchy stress based on the suggestion of the structure of molecular chains can be expressed as [38,41]
\[ \sigma = -p I + \frac{3nk_BT}{b^2} \sum_{n=1}^{N-1} \langle (r_{n+1}(t) - r_n(t))(r_{n+1}(t) - r_n(t)) \rangle \]  
(43)
where $r_n$ is the position vector of the $n$th Kuhn monomers, and $t$ indicates time. From Eq. (43), the stress tensor can be determined once $r_n$ is known. Rouse [64] and Edwards [38] successfully calculated them based on different microscopic pictures.

Rouse [64] regarded $n$ as a continuous index $n$ that runs from 0 to $N$, so that $r_n(t)$ can be transformed as $r(n, t)$ and Eq. (43) can be rewritten as
\[ \sigma = -p I + \frac{3nk_BT}{b^2} \sum_{n=1}^{N} \frac{\langle \partial_t \nabla r \partial_t \nabla r \rangle}{\langle \nabla n \partial n \rangle} \]  
(44)
He further proposed a bead-spring model by representing the single-chain diffusion as Brownian motion, and which led to the expression of $r(n, t)$ as
\[ \frac{\partial r}{\partial t} = \frac{3k_BT}{\xi} \frac{\partial^2 r}{\partial n^2} + \frac{L}{\xi} \cdot r + g(n, t) \]
(45)
\[ g(n, t) = 0; \quad \langle g(n, t) g(m, t') \rangle = 2k_BT \delta_{nm} \delta(t-t') \]

The motion of a molecular chain performing a wormlike random walk in the 'tub' formed by its neighboring chains is frequently referred to as reptation.

\[ \sigma = -p I + \frac{3nk_BT}{b^2} L \sum_{n=0}^{N} \int_0^L \frac{\langle \partial_t \nabla r \partial_t \nabla r \rangle}{\langle \nabla s \partial s \rangle} ds \]  
(47)
Neglecting the compressibility, Cho [41] further reformulated Eq. (47) as
\[ \sigma = -p I + \frac{3nk_BT}{b^2} L \sum_{n=0}^{N} \int_0^L Y(s, t) ds \]  
(48)
where
\[ Y(s, t) \equiv \langle u(s, t)u(s, t) \rangle - \frac{1}{3} I \]  
(49)
with
\[ u(s, t) = \frac{\partial r}{\partial s} \]  
(50)
The dynamic equation can be expressed as [38,41]
\[ \frac{\partial^2 Y}{\partial t^2} = D_c \frac{\partial^2 Y}{\partial s^2} \]  
(51)
with the boundary conditions
\[ Y(0, t) = Y(L, t) = 0 \]  
(52)
where $D_c$ is a diffusion constant.

Equation (49) can be explicitly solved as
\[ Y(s, t) = Z(F) \sum_{m=0}^{\infty} \frac{4}{m!} e^{-t/\lambda_m} \sin \frac{m\pi s}{L}, \quad \lambda_m = \frac{L^2}{m^2 \pi^2 D_c} \]  
(53)
where $F$ is the deformation tensor, $Z(F)$ is given as
\[ Z(F) = \left( \frac{|F \cdot n|}{C : uu} \right) \frac{1}{3} I \]  
(54)
where \( \langle \rangle \) is the ensemble averaging operator. Substituting Eq. (53) into Eq. (48), we have

\[
\sigma = -pI + \frac{3nk_BT}{b^2} \frac{L^2}{N} Z(F)\phi(t)
\]

with

\[
\phi(t) = \sum_{m=odd}^{\infty} \frac{8}{\sigma_m} e^{-t/\sigma_m}
\]

According to Eq. (55), the stress tensor can be determined once the tensorial function \( Z(F) \) is evaluated.

### 3.2 Thermodynamically Based Models

**Lubliner model:** Lubliner [63] extended the work of Green and Tobolsky [66] by introducing the multiplicative decomposition concept into viscoelasticity. He considered three types of rheological models for viscoelasticity under large deformations: Maxwell representation (Fig. 14), Kelvin representation (Fig. 15), and generalized Maxwell representation (Fig. 17). Lubliner [63] assumed the free energy function \( W \) can be decomposed into the volumetric and the distortion parts, and the volumetric deformation is assumed to be purely elastic and the viscoelasticity is caused by the distortion deformation. Lubliner [63] chose the Mooney–Rivlin model as the specific form of the SEDF [48,49] and quantified the evolution laws of internal variables. So the viscohyperelastic model was constructed completely.

**Model of Le Tallec et al.:** Le Tallec et al. [67] developed a viscoelastic framework for incompressible solids based on the Maxwell rheological model (Fig. 14), and the related numerical scheme was proposed for real applications. They presented the framework of a thermodynamic model that incorporated the St.-Venant-Kirchhoff model for the specific SEDF [68].

**Holzapfel-Simo model:** Holzapfel and Simo [69] developed a thermomechanical fully coupled model. This model is based on the concept of internal state variables and is consistent in thermodynamics; i.e., it satisfies Eq. (42). The theoretical frameworks for the SEDF and the evolution equations were summarized, and their model can be interpreted by using the Generalized Maxwell rheological model (Fig. 17). The specific form of the free energy can be chosen from a pool of the SEDF. The generalized St. Venant-Kirchhoff SEDF [68] was used as an example in their paper.

**Reese-Govindjee model:** Based on the work of Lubliner [63], Reese and Govindjee [53] extended the model by considering the finite viscous deformation and nonlinear evolution laws for viscous behavior. Due to the fact that the evolution equation of the internal variables has a similar mathematical structure as the one used in finite elastoplasticity [70], the evolution can be easily implemented in an existing finite element code. Based on the Maxwell rheological model (Fig. 14), whose theoretical framework has been discussed in Sec. 3.1.2 in our paper, the general expressions of the SEDF and evolution equation were formulated in their paper. The specific form of the SEDF in their paper is the Ogden model [50].

**Bonet model:** Bonet [71] decomposed the free energy into volumetric, long-term, and viscous components and derived a new nonlinear internal evolution equation, and the evolution equation was expressed in an incremental form for numerical implementation. Bonet model was derived based on the generalized Maxwell rheological model (Fig. 17) under an isothermal condition. He developed the model for two types of formulations: material formulation and spatial formulation. The material formulation was defined in the reference configuration for describing anisotropic materials, and the related general SEDF and evolution equations were given, respectively. The spatial formulation was employed for isotropic materials, and the related general SEDF and evolution equations were also given, respectively. The specific equations in this paper were presented by using the models of Neo-Hookean [42] and Perić et al. [72].

**Model of Amin et al.:** Amin et al. [59] investigated the rate-dependent behavior of rubbers within compression regimes based on the Maxwell rheological model (Fig. 14). To better characterize the hyperelastic response, a modified hyperelastic model was presented. The evolution equation was given in their paper. Later, Amin et al. [60] continued their efforts, and they investigated the rate-dependent behavior of filled rubbers within compression and shear regimes. They also used the scheme of Maxwell representation (Fig. 14) and derived the general form of the free energy function. Based on the experimental observations, the power laws for evolution equation were proposed in their paper. The specific form of SEDF was given for practical application.

**Hong model:** Hong [55] proposed a theoretical frame to couple the viscoelastic and the electric fields. In this paper, the Maxwell rheological model (Fig. 14) was used, and the general SEDF and evolution equation were derived, respectively. The specific material model was given in this paper as an example.

**Kumar and Lopez-Pamies model:** Kumar and Lopez-Pamies [56] proposed the so-called two potential constitutive framework for rubber viscoelasticity, and the model accounts for the non-Gaussian elasticity of elastomers and the deformation-enhanced shear thinning. Besides, the model was proved to be computationally efficient and robust via comparing it with experimental data of two elastomers. Furthermore, several models, including Le Tallec et al.’s [67], Bergstrom and Boyce’s [21] and Reese and Govindjee’s [53] models, etc., can be regarded as the special cases of their framework. Based on the Maxwell rheological model (Fig. 14), their theoretical framework was constructed with specific forms.

### 3.3 Physically Based Models

**Bergstrom and Boyce model:** Based on a series of experimental data, Bergstrom and Boyce [21] proposed a semi-physically (micromechanism-inspired) based new model with the assumption that the mechanical behavior can be decomposed into an equilibrium network and a nonlinear rate-dependent network which is governed by the reptational motion of molecular chains. The model was developed by using the Maxwell rheological model, as shown in Fig. 14, the spring \( A \) represents the equilibrium network and the spring \( B \) represents the nonlinear rate-dependent network. They characterized the nonlinear rate dependency by assuming the molecular chains doing Brownian motion along the constraint tube.

**Vandoolaeghe and Terentjev model:** Vandoolaeghe and Terentjev [73] extended the classical Rouse model to study the equilibrium and dynamic response of cross-linked polymer within the affine deformation limitation. The general form of stress is formulated in their paper. Merging the Rouse model with the effects of the entanglement, Vandoolaeghe and Terentjev [74] proposed an improved tube model for rubber viscoelasticity but this model did not take into account the non-affine deformation of the molecular network.

**Model of Tang et al.:** Tang et al. [22] proposed a two-scale theory of nonlinear viscoelasticity for unfilled and filled cross-linked elastomers, and the tube model was modified to consider the cross-links by using fractional derivative methods. Besides, the local
inhomogeneous deformation is also described by an extended multi-resolution framework. In their work, the elastomeric microstructure is assumed to consist of the crosslinks and reinforcement superimposed by free chain networks (microzone). For unfilled network, the microzone deforms homogeneously, while the microzone in filled network with weak physical bonds deforms inhomogeneously.

Model of Long et al.: Long et al. [75] proposed a three-dimensional finite strain constitutive model which quantifies the connection between rate-dependent mechanical behavior and kinetics of breaking and reattachment of temporary cross-links in dual cross-linked gels.

Model of Li et al.: Extending the tube theory [46] by considering the deformation of the tube, Li et al. [23] decomposed the network into the hyperelastic cross-linked network and free chains. The viscoelasticity is assumed to be attributed to the diffusion of free chains. Based on the decomposition, they proposed a physically based model to simulate finite strain viscoelasticity, which can be understood by the molecular dynamics method. However, they only considered the disentanglement of free chains and ignored the contour length relaxation.

Model of Xiang et al.: Xiang et al. [24] developed a physically based viscoelastic constitutive model. The stress is decomposed into a hyperelastic part which comes from the elastic ground network (cross-linked network and entanglement network), and a viscous part which is originated from free chains. Utilizing the same scheme from the previous work [19], the free chains are considered to be either: the untangled cross-linked network or entanglement network. The contour length relaxation and disentanglement from the networks of free chains are responsible for the viscous behavior.

3.4 Mixed Models. Miehe-Göktepe model: Miehe and Göktepe [76] proposed a non-affine model for rubber viscoelasticity by adding the contribution of viscous overstress (internal variables) into the equilibrium stress which is from the ground-state network and modeled by their previous work [25]. Their model is thermodynamically consistent, even though the internal variables, formulated physically as the viscous overstress, is attributed to the superimposed chains.

Model of Linder et al.: Following the work of Miehe and Göktepe [76], Linder et al. [77] proposed a new thermodynamically consistent micromechanics-based model. Linder et al. [77] assumed that the viscous stress originates from the transient sub-network that was formed by the temporary entanglements. The microscopic mechanism of their model is equivalent to the generalized Maxwell rheological model (Fig. 17), and the evolution of the sub-network was developed by considering the Brownian motion of the endpoints of the network.

Model of Zhou et al.: More recently, Zhou et al. [78] developed a micro–macro constitutive model that incorporates the nonlinear viscosity, which is related to the diffusion of polymer chains. Their model was constructed within the thermodynamics framework based on the generalized Maxwell representation (Fig. 17); the spring and dashpot in the rheological model are linked to the elastic network and viscous sub-network, respectively. The related viscous stress and internal variables evolution equations were derived by using the modified tube model.

3.5 Summary 2. In Table 4, we summarize the viscohyperelastic models reviewed in Sec. 3. The internal variables and evolution equations, which depend on the materials and loading conditions, are the key characterizations of different models considering the Clausius–Duhem inequality. Here, we classify the models with phenomenological evolution equations as thermodynamically based models, and the ones with evolution equations derived from physical mechanisms as mixed models. Physically based models refer to those derived from the structure of molecular chains without thermodynamic considerations.

4 Damage

4.1 The Basics for the Mullins Effect. The rubbery materials exhibit an obvious degradation in the mechanical behaviors after their first deformation. The change (mainly stress softening) in mechanical properties is named as the Mullins effect due to a series of researches made by Mullins et al. [79,80] and Mullins and Tobin [81,82]. The Mullins effect has been studied for more than one century since the first experimental observation by Bouasse and Carrière [83] in 1903; nevertheless, researchers have not yet reached a unified view of the microscopic mechanism of the Mullins effect.

The Mullins effect was mostly discovered in particle-filled rubber-like materials. Buehe attributed the Mullins softening effect to tearing molecular chains off the surface of particles as well as breaking of chain between filler particles [84,85]. Later on, Harwood et al. and Harwood and Payne reported that the unfilled pure gum and unfilled vulcanizate show stress softening phenomenon as well [86,87]. They pointed out that the possible sources of the Mullins effect on unfilled rubbers include the following: (1) breaking of crosslinks, (2) breaking of network chains and (3) residual local orientation of polymer chains.

The breaking of covalent bonds of crosslinks and chains causes permanent damage showing contrary to the recovery of the Mullins softening effect. Hence, a chain slipping model was proposed by Houwink to explain this phenomenon [88]. The polymer chain slips along the surfaces of filler particles when the deformation of reinforced rubber increases and the reversible bonds between chain and fillers break at the same time. After unloading, new bonds reform between polymer chain and fillers, resulting in a longer polymer chain between fillers (because slipping does not work in unloading), and the chains start to slip again once the historical maximal deformation is exceeded. The degradation of materials (entropy change) can be restored by increasing the temperature. This microscopic physical picture interprets the stress softening and restoration well. The discussion earlier does not involve the irrecoverable breaking of covalent bonds between polymer chains and fillers.

Kraus et al. [89] showed the small change of network chain density by the swelling experiment of vulcanizates. Meanwhile, the inappreciable volume expansion of particle-filled styrene-butadiene rubber (SBR) confirmed that few bonds break when stretched to 300% elongation. Therefore, Kraus et al. [89] concluded that bond breaking is not the only reason for the Mullins effect. They believed that the breaking of filler structure combined with the breaking of the bonds accounts for the Mullins effect. Moreover, the effect of filler rapture appears to be more essential in the condition of lower temperature, higher strain rate, and higher filler concentration.

Hanson et al. [90] proposed a new microscopic explanation for the Mullins effect and attributed the stress softening to the disentanglement. They observed that the Mullins effect disappears for the silica-filled polydimethylsiloxane (PDMS) when the second stretch direction is perpendicular to the first stretch direction.
According to the physical picture they proposed, once the silica-filled PDMS is deformed, one polymer chain slides through the other chain at its attachment point to the filler particle, thus removing the entanglement. The irreversible and directional removal of entanglement leads to the anisotropic Mullins effect. Furthermore, the cross-linked chain density keeps the same while the entanglement density decreases, this exactly is the most prominent difference from other micro interpretations.

The aforementioned physical mechanisms for the Mullins effect are based on the observations of mechanical experiments. Researchers also carried out other experiments to verify their microscopic interpretations. Suzuki et al. [91] measured the chain scission of silica-filled SBRs whose interfacial properties between polymer matrix and fillers vary. The carbon radicals formed by chain scission can be accurately detected by the electron spin resonance (ESR) measurements. The experiments showed that the stronger the interfacial bonding between polymer and fillers is, the larger the increase of carbon radicals is, and the more prominent the Mullins effect is. Thus, Suzuki et al. concluded that chain scission might contribute to the Mullins effect. Later, Ducrot et al. [92] introduced a special chemoluminescent cross-linker (which emits light when breaks) into elastomers. The light emission helps to indicate where bond breakage happens in real-time, and it shows the potential to research the origin of the Mullins effect. Following the study of Ducrot et al. [92], recent research by Clough et al. [93] demonstrated that the scission of even a small quantity (<0.1%) of covalent bonds contributes distinctly to the Mullins effect of silica-filled PDMS. Furthermore, they showed unambiguously that covalent bond scission happens in an anisotropic way, which results in the anisotropy of the Mullins effect.

With the understanding of the microscopic mechanism behind the Mullins effect, researchers developed mechanical models, including the physically based models, thermodynamically based models, as well as the mixed models, to quantitatively describe the Mullins effect. Here, we introduce models for the Mullins effect as comprehensively as possible, and some recently proposed and representative models will be analyzed in detail in later chapters.

Based on the assumption that the particle reinforced rubber consists of soft and hard regions, Mullins and Tobin [81] put forward a phenomenological two-phase model to characterize the mechanical behaviors of the filler reinforced rubbers. Since the deformation of the hard region is assumed to be negligible, the overall strain of rubber is proportional to the strain of the soft region. The hard region can be transformed into a soft region under large deformation, and the volume fraction of the soft region increases as the historical maximal stress increases. Followed by this work, Mullins and Tobin [82] proposed a strain amplification factor to quantify the relationship between the strain of the soft region and the average strain of the rubber. Johnson and Beatty [94] carried out an extended exploration of the two-phase theory and strain amplification factor to quantify stress softening, and their model was applied to a dynamic problem for the first time and then applied to capture the Mullins effect in equibiaxial extension (inflation of a balloon) [95]. Thereafter, the strain amplification expression for uniaxial tension was extended to a general three-dimensional deformation state [96] and was further incorporated into a constitutive model for Mullins effect [97] (more details in Sec. 4.4).

The notion by Bueche [84], i.e., stretch-induced debonding between filler and polymer matrix, was employed by Govindjee and Simo [98] to form a physically based continuum damage model. The feature of this model is the decomposition of free energy into the contribution from the chain network between cross-linkers and from chain network between filler particles. Soon after, this micro-mechanical model was transformed into a phenomenological one to improve computing efficiency [99]. Then, Göktepe and Miehe [100] extended the isotropic theory of Govindjee and Simo [99] into an anisotropic model. Lion [101] developed a macroscopic theory for filled rubber, among which the Mullins effect was described by a continuum damage model, and the evolution of damage depends on the historical maximal strain. There are some other macroscopic models based on the continuum damage mechanics, and the classic examples include Refs. [102–106].

The residual deformation and anisotropy of the Mullins effect were taken into account in some models. Based on the theory of pseudo-elasticity proposed by Ogden and Roxburgh [107] (more details in Sec. 4.2), Dorfmann and Ogden [108] developed a phenomenological constitutive model for Mullins effect with residual deformation. Two internal variables were adopted into the strain energy function to separately capture the stress softening and residual deformation. Based on the physical mechanism of the Mullins effect, Göktepe and Miehe [100] constructed an anisotropic Mullins-type damage model grounded on the micro-sphere model (more details in Sec. 4.4). Molecular chains along different directions have different elongations and different degrees of damage, leading to the residual deformation and anisotropy. The works based on the micro-sphere model includes Refs. [109–111]. Recently, Zhong et al. [112] took the micro-sphere model to describe the damage of cross-linked network and considered the degradation of entanglement (more details in Sec. 4.3).

### 4.2 Thermodynamically Based Models. Ogden-Roxburgh model: Ogden and Roxburgh [107] developed a pseudo-elastic model for the Mullins effect of rubbery material. Considering a specification of the model to a biaxial deformation

\[
\eta = 1 - \frac{1}{r} \text{erf} \left[ \frac{1}{\eta m} (W_m - W(\lambda_1, \lambda_2)) \right]
\]  

\[
W_m = W(\lambda_{1m}, \lambda_{2m})
\]

\[
\sigma_\beta - \sigma_3 = \lambda_\beta \frac{\partial W}{\partial \lambda_\beta} = \eta \sigma_\beta \frac{\partial W}{\partial \lambda_3} = \eta (\sigma_\beta - \sigma_3)
\]  

\[\beta = 1, 2 \] (59)

Here, (\cdot) denotes the quantities without damage effect. For example, \(W\) is the SEDF of the purely hyperelastic material. \(W\) is the corresponding SEDF considering the damage effect. The damage parameter \(\eta\) connects the stress after damage and its corresponding perfectly elastic stress (Eq. (59)). \(\eta\) is expressed as an error function of \(W\) and its corresponding value \(W_m\) defined by Eq. (58), with \(\lambda_{1m}\) being the value of \(\lambda_1, \lambda_2\) at the point at which unloading begins. The parameter \(r\) represents the degree of damage, and \(m\) describes the dependence of damage on deformation.

When the material deforms along a primary loading path for the first time, the damage parameter keeps a constant \(\eta = 1\), which means the material behaves like an intact hyperelastic material. The unloading from any point on the primary loading path activates the damage parameter. For the first unloading and the subsequent submaximal reloading and unloading, \(\eta\) develops in accordance with Eq. (57) \((0 < \eta < 1)\). Once \(W(\lambda_1, \lambda_2)\) reaches \(W_m\), we regain \(\eta = 1\). And the loading curve rejoins the primary loading path, with \(W_m\) updated to a new value.

According to Eq. (57), the damage parameter depends on the energy rather than the deformation or stretch, which makes this model different from the models whose damage criteria are based on deformation (e.g., the network alteration model [26]). Any pairs of \((\lambda_1, \lambda_2)\) that satisfies \(W_m = W(\lambda_1, \lambda_2)\) can be regarded as a starting point for new damage. Similarly, to quantitatively describe the pseudo-elasticity of double-network hydrogel, Wang and Hong [112] built a relationship between stiffness of double-network hydrogel and its corresponding value after damage by parameter \(\eta\), which was expressed as a function of historical maximal stretch. Lu et al. [113] adopted an analogous form of the parameter \(\eta\) to build a constitutive model incorporating the Mullins effect for soft materials. By adopting the softening variable \(\eta\), Wang et al. [114] developed a phenomenological model which captures the Mullins effect and the shakedown phenomenon of tough hydrogels under cyclic loads. Recently, Lu et al. [115] put forward another pseudo-
elasticty theory by introducing an internal variable into the strain energy function of a single chain. This theory is capable of modeling the Mullins effect and the complex rate-dependent behaviors of tough hydrogels.

In this model, the dissipation rate is non-negative, which means Clausius–Duhem inequality is satisfied here. This thermodynamically based model can reasonably capture the damage behaviors of rubbery materials, while the damage parameter setting is inconsistent with the actual damage development in rubbery materials to some extent. This model is phenomenological.

4.3 Physically Based Models. Model of Marckmann et al.: Marckmann et al. [26] developed a new network alteration theory to describe the Mullins effect. For both the particle-filled elastomer and none particle-filled elastomer, chain fracture and linker breakage happen as the applied deformation increases. Consequently, the number density of chains decreases and the number of Kuhn monomers of a single chain (the average chain length) increases, which are described by the following three equations:

\[ N = N(\lambda_{\text{max}}) \]  
\[ n = n(\lambda_{\text{max}}) \]  
\[ n \cdot N = \text{constant} \]  

where \( N \) denotes the average chain length and \( n \) is the number density of chains. The product of \( N \times n \) signifies the number of monomers per unit volume which keeps a constant. While the details of the network alteration (fracture of polymer chains and linkers) can hardly be observed experimentally, the specific functional form of \( N \) and \( n \) depend on the mechanical responses of specific materials, and the exponential function and polynomial expression are mostly used [26,109]. By adopting the eight-chain model, the Cauchy stress can be expressed as

\[ \sigma_i = -p + \frac{1}{3} C_k(\lambda_{\text{max}}) \sqrt{N(\lambda_{\text{max}}) \lambda_i^2} L^{-1} \left( \frac{\lambda_i}{\sqrt{N(\lambda_{\text{max}})}} \right) \]  

\[ i = 1, 2, 3 \]  

with

\[ C_k(\lambda_{\text{max}}) = n(\lambda_{\text{max}}) k_B T \]  
\[ \lambda_i = \sqrt{I_i(t)/3} \]  
\[ \lambda_{\text{max}} = \max_{0 \leq \tau \leq \tau^*} [\lambda(\tau)] \]  

where \( p \) is the hydrostatic pressure which can be determined by the boundary condition, \( C_k \) is the modulus, \( \lambda_i \) is the principal stretch, and \( \lambda_{\text{max}} \) is the historical maximal chain stretch.

The network alteration theory was further improved by Chagnon et al. [116] where Eq. (62) is no longer valid because chain breakage leads to the formation of dangling chains. Two equations are required to describe the alteration of chain length \( N \) and chain density \( n \), respectively.

The network alteration theory is grounded on the molecular chain microstructure. Both the physical meaning and mathematical expression are explicit. This model can be further improved to describe the damage-induced anisotropy and residual deformation. In the Ogden-Roxburgh model, during the first loading, the damage parameter is inoperative, until unloading occurs; on the contrary, in the Model of Marckmann et al., damage occurs immediately during the first loading, and the damage-related variables no longer develop in the subsequent unloading. The two models have completely opposite settings for damage variables, while the latter is closer to reality.

Model of Zhong et al.: Recently, Zhong et al. [27] proposed a damage model for soft materials. The SEDF is composed of the cross-linked part and the entangled part, as discussed in Sec. 2.2.2 of this paper. For the cross-linked part, they adopted the network alteration theory to describe the damage effect, and they used the micro-sphere model to incorporate the damage-induced anisotropy and residual deformation. For the entangled part, as the deformation increases, the surrounding chains break and the entangled constraint acting on an individual chain will decrease. The irreversible degradation of entangled constraint is reflected by a decreasing entangled modulus. The principal stress along direction \( i \) with damage effect is expressed as

\[ \sigma_i = \sum_j \text{weight}(j) \frac{3G_c \lambda_i^2 (\alpha_j)^2}{(1 - (\lambda_{\text{chain}}^0)N^j)(1 + (\lambda_{\text{chain}}^0)^2/2N^j)} \]

\[ -G_c \frac{1}{k_i} + p_{\text{hydro}} \]  

(67)

Here, \( i, j = 1, 2, 3 \) indicates three principal directions and \( j = 1, 2, \ldots, J \) stands for the chain series number. The symbols \( \alpha_j \) \( i = 1, 2, 3 \) = \([\alpha', \beta', \gamma']\) denote the end-to-end unit vector of the chains along the \( j \)th direction, and \( weight(j) \) is the chain density weight along the \( j \)th direction, and \( p_{\text{hydro}} \) is the hydrostatic pressure. The cross-linked modulus for the \( j \)th direction is expressed as

\[ G_c^j = G_{c0}/\text{DamageCo}(t) \]  

with the initial (non-damage) cross-linked modulus \( G_{c0} \). The number of monomers of the chains along \( j \)th direction is expressed as

\[ N^j = N_0 \cdot \text{DamageCo}(t) \]  

(69)

with \( N_0 \) being the initial number of monomers per chain. The damage coefficient for the \( j \)th direction chains is

\[ \text{DamageCo}(t) = \max_{0 \leq \tau \leq \tau^*} [p \cdot (\lambda_{\text{chain}}^j(t) - 1)^2 + 1] \]  

(70)

with \( p \) being the damage factor of the cross-linked network, and the stretch of the chains along the \( j \)th direction given by

\[ \lambda_{\text{chain}}^j = \sqrt{\lambda_1^j \alpha_1^j + \lambda_2^j \beta_1^j + \lambda_3^j \gamma_1^j} \]  

(71)

The entangled modulus is expressed as

\[ G_c = \frac{G_{c0}}{\sqrt{\exp(k(\lambda_{\text{max}} - 1))}} \]  

(72)

\[ \Lambda_{\text{max}} = \max_{0 \leq \tau \leq \tau^*} [\sqrt{I_i(t)/3}] \]  

(73)

with the initial entangled modulus \( G_{c0} \), the damage factor of the entangled network \( k \), and the historical maximal value of macroscopic deformation \( \Lambda_{\text{max}} \). For the cross-linked network, they distinguished the various chain stretches along different directions to incorporate the damage-induced anisotropy and adopted a unified macroscopic stretch for the entangled network to avoid complexity.

Zhao model: Zhao [28] developed a theory to characterize the damage of interpenetrating polymer networks (IPN). Taking the double-network hydrogel [117] as an example. This IPN includes a highly cross-linked network \( A \) and a loosely cross-linked network \( B \). The interpenetration of network \( A \) stretches the polymer chains of network \( B \) and reduces its chain density. At the same time, the chains of network \( A \) are also stretched by network
\[ B \text{ accompanied by a decreased chain density. Pre-stretch exists for the two networks and can be expressed as } C_i^{-1/3}, \quad \text{with} \]

\[ C_i = \frac{n_i^2}{n_i} \quad i = A, B \quad (74) \]

where \( n_i \) is the number density of chains of the \( i \)th network without pre-stretch and \( n_i^2 \) is the number density of chains of the \( i \)th network of the IPN. Due to the existence of pre-stretch, the total chain stretch can be expressed as

\[ \lambda_i = C_i^{-1/3}, \quad \lambda_i' = C_i^{-1/3} \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2} \quad (75) \]

where \( \lambda_i' \) corresponds to the chain stretch because of the deformation of IPN caused by the external loading.

Network \( A \) is highly cross-linked so its chain extension limit is much smaller compared with network \( B \). The damage of network \( B \) is ignored and the damage of network \( A \) is described by the network alteration theory

\[ N_i = N_{A0} \exp \left[ q(A_{A0}^{\text{max}} - C_i^{-1/3}) \right] \quad (76) \]

\[ n_i = n_{A0} \exp \left[ -p(A_{A0}^{\text{max}} - C_i^{-1/3}) \right] \quad (77) \]

where \( n_{A0} \) is the initial chain density of network \( A \), and \( N_{A0} \) is the original number of freely joint links of an individual chain in network \( A \), the symbols \( p \) and \( q \) denote material constants. \( A_{A0}^{\text{max}} \) is the historical maximal total chain stretch of network \( A \).

Based on the above assumptions, the uniaxial tension/compression stress can be expressed as

\[ \sigma_j = \left( \frac{C_A^{1/3} \zeta}{3n_A n_A} \right) \exp \left[ \left( -\frac{1}{2} q - p \right) \left( A_{A0}^{\text{max}} - C_A^{-1/3} \right) \right] + C_B^{1/3} \zeta \frac{1}{3n_A n_A} \left( \lambda_1^2 - 1 \right) \lambda_i \quad (78) \]

where \( \zeta \) and \( \lambda_i \) are the volumes of one monomer of network \( A \) and \( B \), respectively, and \( \lambda_i \) is the inverse Landegevin function.

By integrating the network alteration theory and the interpenetrating network theory, Zhao model predicts the Mullins effect of IPN and explains the necking instability phenomenon observed in the experiments of double-network hydrogel.

Zhu-Zhong Model: Zhu and Zhong [118] proposed a model for double-network hydrogel. Similar to the Zhao model, they split the double-network hydrogel into network \( A \) and network \( B \), with network \( A \) damaged after deformation and network \( B \) fully elastic. The pre-stretches and total chain stretches of these two networks are described by Eqs. (74) and (75).

Zhu and Zhong [118] took the chain length \( n_{A} \) and chain density \( n_i \) of the network \( A \) as two internal variables to quantify the damage evolution. However, different from Zhao model which adopted empirical equations (Eqs. (76) and (77)), these two parameters are determined from analyzing the energy dissipation. First, similar to the energy dissipation by the viscoelasticity as presented in Section 3.1.1, the energy dissipated by internal fracture of network \( A \) satisfies the Clausius-Duhem inequality (Eqs. (34) and (36)), which is expressed by a specific form here as

\[ \dot{\varepsilon} = \left[ \frac{\partial W}{\partial N_A} \xi_T + \frac{\partial W}{\partial n_A} \right] \geq 0 \quad (79) \]

Based on the assumption that no dangling chain is produced during the damage of network \( A \), the product of \( N_A \) and \( n_A \) keeps a constant, set as \( \Phi_A \)

\[ N_A \cdot n_A = N_{A0} \cdot n_{A0} = \Phi_A \quad (80) \]

with \( N_{A0} \) and \( n_{A0} \) are the corresponding quantities of \( N_A \) and \( n_A \) before any damage happens. Then, the energy dissipation rate can be expressed in an incremental form

\[ \dot{\varepsilon} = \left[ \frac{\partial W}{\partial N_A} \Phi_A \frac{\partial W}{\partial n_A} + \frac{\partial W}{\partial n_A} \right] \Phi_A (\varepsilon_{A0}) \quad \text{if } \varepsilon = \varepsilon_{A0} \geq \varepsilon_{\text{critical}} \quad (81) \]

It is obvious from Eq. (81) that the evolution of the energy dissipation with respect to deformation is related to the development of the parameter \( N_A \). When the deformation of double-network hydrogel is smaller than its historical maximal value, no additional energy is dissipated and the parameter \( N_A \) remains unchanged. Once the deformation exceeds its historical maximal value which is larger than a critical strain, the damage further develops and part of the mechanical energy is dissipated. Once we get the specific form of the energy dissipation \( \dot{\varepsilon} \) based on the experimental data (e.g., the hysteresis loop between loading and unloading curves), the damage evolution law, i.e., the evolution law for the two internal variables can be solved. And the constitutive formulation of the double-network elastomer is determined.

In the models based on network alteration theory, like the Model of Marxmann et al., the Model of Zhong et al. and Zhao Model, as we mentioned above, damage happens immediately once the external force is applied. Based on the experimental observations of the double-network hydrogel [119], Zhu and Zhong set a critical strain, above which the damage initiates. In the Zhao Model, the evolution laws of the chain length and chain density are assumed in advance (Eqs. (76) and (77)), and the damage parameters are determined by mathematically fitting the theoretical stress–stretch relationship with experimental data, like the overall loading/unloading stress–stretch data. While the development of the chain length and chain density in this model is decided by analyzing the energy dissipation (Eq. (81)) and mass conservation (Eq. (80)), several sets of data on both stretch value and area of hysteresis loop determine the energy dissipation function \( \dot{\varepsilon} \). In terms of model applicability, the Zhu-Zhong Model captures the mechanical behavior of double-network hydrogel before the necking, while the Zhao Model further describes the hardening phenomenon after the necking of double-network hydrogel.

**Model of Lavoie et al.:** Lavoie et al. [29] developed a continuum model to describe the progressive damage of multi-network elastomer (MNE). MNE composed of \( M \) networks, and the deformation gradient of the \( i \)th network can be expressed as

\[ F_M = \textbf{F}_M \phi_{M-1} \quad (82) \]

and the corresponding left Cauchy–Green deformation tensor is

\[ B_M = F_M^T (F_M)^T \quad (83) \]

where \( \textbf{F} \) corresponds to the deformation applied on the MNE by the external loading, and \( \phi_{M-1} \) is the deformation gradient for the \( i \)th network caused by the \( M-i \) times swelling and drying operations.

The total SEDF of MNE is the sum of each network’s energy density times its volume fraction \( \phi_M \) as

\[ W_M(B) = \sum_{i=1}^{M} \phi_M W_M(B_M) \quad (84) \]

and the Cauchy stress is

\[ \sigma_M = \sum_{i=1}^{M} \phi_M^2 \frac{\partial W_M(B_M)}{\partial B_M} - p I \quad (85) \]

Similar to the double-network hydrogel, the mechanical response of the filled network (\( i = 1 \)) varies from the matrix networks (\( i > 1 \)), so we need different SEDFs to describe them.

The SEDF for the matrix network is expressed by the generalized neo-Hookean constitutive model as

\[ W_M(B_M) = \frac{\mu_M}{2n_M} \left[ (1 + (L_i B_M)^3) - 3 \right], \quad i > 1 \quad (86) \]

with a shear modulus \( \mu_M \) and a material parameter \( n_M \).
The SEDF for the filled network can be expressed as
\[ W_M^1 = \rho \int f(N)B(r_{\text{max}}^*, N)E_{ch}(r^*)dN \]  
(87)
where \( \mu = nk_BT \), \( n \) is chain density without swelling or drying, \( f(N) \) is the probability density function that describes the distribution of chain lengths, \( B(r_{\text{max}}^*, N) \) is the damage function which depends on both the historical maximal chain stretch \( r_{\text{max}}^* \) and the number of Kuhn monomer per chain \( N \). \( E_{ch}(r^*) \) is the dimensionless free energy of the stretched polymer chain, and \( r^* \) is the fractional stretch of polymer chains.

The probability density function \( f(N) \) can be determined by some experimental methods. For example, Ducrot et al. [92] synthesized polymers with specific cross-linkers, and these cross-linkers emit light when they break. The recorded light intensity can estimate the chain length distribution. The Maxwell–Boltzmann distribution can also be adopted, and this distribution was incorporated into a rate-dependent model to describe the progressive damage behavior of elastomers [120]. The damage function \( B(r_{\text{max}}^*, N) \) depends on both the historical maximal chain stretch \( r_{\text{max}}^* \) and the number of Kuhn monomer per chain \( N \). The interested readers may refer to Ref. [29] for more details about the damage function. The fractional stretch \( r^* = R/Nb \) of polymer chain is given as
\[ r^* = \frac{R}{Nb} = \sqrt{\frac{L(B)}{3N}} \]  
(88)
where \( R \) is the end-to-end distance of a single chain and \( b \) is the length of a Kuhn monomer. The dimensionless polymer chain energy \( E_{ch}(r^*) \) is expressed as
\[ E_{ch}(r^*) = \int_0^{r^*} F_{ch}(r^*)dr^* \]  
(89)
where \( r_0 = \sqrt{1/3N} \) is the initial fractional chain stretch. The dimensionless chain force is expressed as
\[ F_{ch} = \begin{cases} \frac{1}{2}(1 - r^*)^2 - \frac{1}{2} - 2r^* & r^* < 0.9 \\ \frac{513}{20} + 501(r^*-0.9) + 26.238(r^*-0.9)^2 \\ + 68.436(r^*-0.9)^3 & r^* \geq 0.9 \end{cases} \]  
(90)
In this model, the distribution of chain length is taken into account. The molecular chain with a specific length has a related damage function. A new force–stretch relationship for a single chain is adopted for avoiding the singularity of chain force when \( r^* \) gets close to 1. This model provides an accurate description of the experimental data of multi-network elastomers.

4.4 Mixed Models. Qi-Boyce model: Following the concept of soft/hard domain by Mullins and Tobin [81], Qi and Boyce [97] developed another damage model. The macroscopic deformation and the deformation of the soft domain are connected by the amplification factor, and this factor is affected by the proportion of the soft domain. As the historical maximal deformation increases, the volume fraction of the soft domain increases. The SEDF \( W \) of the material is assumed to be only contributed by the soft domain and is expressed as follows:
\[ W = v_Nnk_BT \sqrt{N} \Lambda_{\text{chain}} \beta + N \ln \left( \frac{\beta}{\sinh \beta} \right) \]  
(91)
where \( \beta \) is the inverse of Langevin function formulated as
\[ \beta = L^{-1}(\Lambda_{\text{chain}}/\sqrt{N}) \]  
(92)

The chain stretch of soft domain \( \Lambda_{\text{chain}} \) is formulated by the macroscopic deformation \( I_1 \) and the amplification factor \( X \) as
\[ \Lambda_{\text{chain}} = \sqrt{X(I_1/3 - 1) + 1} \]  
(93)
and the amplification factor \( X \) is expressed as
\[ X = 1 + 3.5(1 - v_s) + 18(1 - v_s)^2 \]  
(94)
with \( v_s \) being the volume fraction of the soft domain. The quantity \( v_s \) varies with the historical maximal deformation and is described as
\[ \dot{v}_s = A(v_s - v_s^c) \frac{\sqrt{N} - 1}{N - \Lambda_{\text{chain}}^0 \Lambda_{\text{chain}}} \]  
(95)
where \( A \) and \( v_s^c \) are material constants.

Once the SEDF \( W \) is obtained, the Cauchy stress \( \sigma \) can be calculated by
\[ \sigma = \frac{v_Nnk_BT}{3} \frac{\sqrt{N}}{\Lambda_{\text{chain}}} L^{-1} \left( \frac{\Lambda_{\text{chain}}}{\sqrt{N}} \right) B - \rho I \]  
(97)
with the left Cauchy–Green deformation tensor \( B \). The symbol \( \rho \) denotes the hydrostatic pressure that can be determined by boundary conditions, and \( I \) is the identity tensor.

This model adopted the concept of soft and hard phases and the amplification factor. It well predicts the mechanical behaviors during cyclic loading under various states of deformation. Giöktepe-Miche model: In order to describe the anisotropic damage behavior of particle-filled rubber-like materials, Güloktepe and Miehe [100] extended their micro-sphere hyperelastic model for considering the damage. The overall network consists of the hyperelastic cross-link to cross-link (CC) part and the particle-to-particle (PP) part which is assumed to be the source of the damage. The mechanical response of the hyperelastic CC network is described by the micro-sphere model [25], which has been introduced in Sec. 2.2.2 of this review. Attention herein is focused on the discussion of the PP network. The damage behavior is originated in the PP network due to the destruction of bonds between polymer chains and filled particles. The main equation in this model is
\[ \beta_{pp}^* = Z(\phi, d) \cdot \beta_{pp}^0 \]  
(98)
where \( \phi \) and \( d \) are the normalized stress function and the normalized energy of single-chain, respectively, \( d \) is the internal damage variable, \( \beta_{pp}^* \) is the micro stress after damage, and \( \beta_{pp}^0 \) is the corresponding perfectly elastic micro stress as
\[ \beta_{pp}^0 = kTN_{pp}^0 \phi(\lambda) \]  
(99)
with \( N_{pp}^0 \) being the number of chain segments (Kuhn monomers) of the PP network. The \( \phi(\lambda) \) is the derivative of \( \phi \) with respect to \( \lambda \) and the expression of \( \phi \) is
\[ \phi(\lambda) = \lambda_{pp}^{\beta} \beta(\lambda_{pp}^{\beta}) + \ln \frac{\beta(\lambda_{pp}^{\beta})}{\sinh \beta(\lambda_{pp}^{\beta})} \]  
(100)
where \( \beta \) is the inverse of Langevin function, the relative stretch \( \lambda_{pp}^{\beta} \) is written as
\[ \lambda_{pp}^{\beta} = \sqrt[N_{pp}^0]{\lambda} \]  
(101)
here \( \lambda \) in Eq. (101) stands for the micro stretch of a single chain, and the micro stretch of a single chain is connected to the macro stretch by the affine relationship shown as
\[ \lambda = \frac{\lambda_{pp}^{\beta}}{\sqrt[N_{pp}^0]{\lambda}} \]  
(102)
with the macro stretch $λ$ being given by

$$λ = √g t ⋅ I$$  (103)

and

$$t = Fr$$  (104)

where $g$ is the Kronecker symbol of the current configuration and $r$ is a unit vector. $F$ is the macroscopic deformation gradient tensor.

As shown in Eq. (98), the perfectly elastic micro stress $β_{pp}$ and the corresponding micro stress after damage $β_{pp}$ are connected by the normalized stress function $ξ$. The specific expression of $ξ$ is given as

$$ξ(a, d) = c_1(d)|a|^3 + c_2(d)$$  (105)

$$c_1(d) = k_a \exp((-1)^n_1 n_0 d) \quad a = 1, 2, 3$$  (106)

$$d = \begin{cases} \phi & \text{for } a = a(d) \text{ and } \phi > 0 \\ 0 & \text{otherwise} \end{cases}$$  (107)

where $[k_a, \nu_1], a = 1, 2, 3$ are six material parameters. The relationship between internal damage variable $d$ and the normalized energy $φ$ is expressed as Eq. (107).

Once the micro stress after damage $β_{pp}$ is obtained, the macroscopic Kirchhoff stress and the tangent modulus can be calculated by

$$\mathcal{F} = \left\{ n_{pp} β_{pp} λ^{-\frac{1}{2}} t \otimes t \right\}$$  (108)

$$θ = 2n_3 n_{pp}$$  (109)

where $n_{pp}$ is the chain density of the PP network. The average operator "<" is in Eq. (108) is interpreted as the homogenization over the surface of a unit micro-sphere [100].

In this model, the micro-sphere model is adopted to characterize the residual deformation and damage-induced anisotropy. There are 13 parameters that need to be fitted by experimental data in total, in which five parameters are from the CC network and others are from the PP network.

Model of Vernerey et al.: Vernerey et al. [121] developed a statistical damage model by analyzing the chain configuration. For a polymer chain consisting of $N$ Kuhn monomers, the end-to-end vector $R$ satisfies the following distribution

$$Φ_0(R) = c_0 \left( \frac{3}{2π N β} \right)^{3/2} \exp \left( -\frac{3 \mathbf{R} \cdot \mathbf{R}}{2 N β} \right)$$  (110)

where $R$ is the undeformed end-to-end vector of a single chain, $c_0$ the initial total chain density, and $Φ_0$ is the chain distribution of undamaged state. The chain distribution can also be expressed in the current state as $Φ_0(R, t)$, where $R$ is the current end-to-end vector of a single chain. When the end-to-end distance of a single chain exceeds a critical value, the chain rapture, and the breakage of the chain then causes damage to the polymer network; as a result, the chain distribution changes as

$$Φ(R, t) = Φ_0(R)(1 - Δ(R, t))$$  (111)

or expressed in the current configuration as

$$Φ(R, t) = Φ_0(R)(1 - δ(R, t))$$  (112)

where $Φ$ is the current end-to-end distance distribution and $Φ$ is the end-to-end distance distribution of undeformed configuration. The quantities with subscript "0" correspond to the non-damage situation. The notation $Δ(δ)$ describes the dimensionless damage distribution which describes the proportion of fractured chains. $Δ(δ)$ is a distribution function about chain length and it evolves with time (deformation). The most essential thing is to find the evolution law of the damage distribution. It is not difficult to imagine that the larger the initial end-to-end distance of a single chain is, the easier it breaks. Vernerey et al. [121] adopted a cumulative probability function $P(r)$ to solve the damage distribution

$$P(r) = \frac{1}{√2 π r_d^2} \exp \left( -\frac{1}{2 r_d^2} (ξ - r_c)^2 \right) dξ$$  (113)

where $r_d$ stands for the standard deviation and it describes the variation in chain fracture, $r_c$ is a critical fracture end-to-end distance of a chain. A basic feature for $P(r)$ is that it is close to zero when the end-to-end distance is small, while increases sharply to one when the end-to-end distance approaches $r_c$. Accordingly, the damage evolution equation can be phrased as

$$Δ(R, t) = ∇P(r) \cdot t ∇P(r) \cdot t ≥ 0$$  (114)

$$Δ(R, t) = 0 ∇P(r) \cdot t < 0$$  (115)

with the initial value of $Δ$ as

$$Δ(R, 0) = P(r)$$  (116)

Once the evolution of damage is confirmed, the chain distribution at any current configuration can be calculated, and the Cauchy stress is

$$σ = ∫_Ω (φ - Φ)(t \otimes r) dΩ + pI$$  (117)

where the integral is taken over all configuration space $Ω$. The quantity $p$ is determined by the boundary condition and $I$ is the identity tensor. $t = dψ/dr$ is the chain force vector, with the chain energy $ψ$ expressed as

$$ψ(r) = k_b T N \left( \frac{λ}{√N} β + \ln \frac{β}{β} \right)$$  (118)

This model describes the end-to-end distance distribution of a single chain quantitatively. Once the end-to-end distance exceeds a critical value, the chain fractures and thus causes damage. The highlights of this model include the clear microscopic physical meaning as well as the capability to accurately characterize the macroscopic mechanical response.

4.5 Summary 3. In Table 5, we summarize the damage models reviewed in Sec. 4. All the models introduced in this section have a damage criterion. The core problem for building a damage model is to construct reasonable damage criteria. For physically based models, all of their damage variables bear physical meanings. If the damage variables do not have physical meaning but the models nevertheless satisfy the Clausius–Duhem inequality, we call them thermodynamically based models. If parts of their damage variables have physical meaning, we term them as mixed models.

<table>
<thead>
<tr>
<th>Damage model</th>
<th>Thermodynamically based models</th>
<th>Mixed models</th>
<th>Physically based models</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Ogden and Roxburgh, 1999)</td>
<td>(Qi and Boyce, 2004); (Göktepe and Miehe, 2005); (Vernerey et al., 2018)</td>
<td>(Markmann et al., 2002); (Zhao, 2012); (Zhang et al., 2019); (Lavoie et al., 2019)</td>
<td>(Zhu and Zhong, 2020)</td>
</tr>
</tbody>
</table>
5 Conclusions and Remarks

This paper presents an overview of constitutive models accounting for the hyperelasticity, visco-hyperelasticity, and damage phenomena in soft materials. We specifically focus on the so-called physically based models accounting for microscopic polymer chain characteristics such as, for example, polymer chain entanglement. Typically, the models that take into account the entanglement constraints provide more accurate fitting of the experimental data. On the other hand, one may want to limit the number of fitting parameters; in this case, sufficient accuracy still can be achieved, as, for example, for Arruda–Boyce model [44] with only two parameters. However, when higher precision is required, especially, under large non-trivial deformation modes, the models, considering entanglement constraints, are recommended. Among such models, the model by Xiang et al. [19] and the non-affine network model [20] show the best capability to fit the experimental data under large deformation. An attractive feature of these accurate models is that they require only three parameters that are directly related to the physical parameters at the microscopic length scales.

The key to developing thermodynamically based viscohyperelastic models resides in the choice of the SEDF, internal variables, and their evolution equations. Physically based models derived from the motion of molecular chains provide a better understanding of the viscosity origin. Quantitatively comparing the performance of various viscohyperelastic models is, however, challenging due to their different application conditions. Nevertheless, we point out some characteristics for these models to help choose suitable models for particular material systems. Holzapfel model [69] can be used to deal with the fully thermomechanical coupled (thermoviscoelasticity) problems. Bonet model [71] is applicable to the anisotropic material. The model of Amin et al. [59,60] is specially developed for high damping rubbers under compression and shear regimes. Kumar and Lopez-Pamies model [56] can characterize the deformation-enhanced shear thinning behavior. The model of Liu and Boyce model [22] can describe the local inhomogeneous deformation for both unfilled and filled cross-linked elastomers. The model of Long et al. [75] aims at the dual cross-link self-healing gel. The model of Li et al. [23] directly derives the material parameters from molecular dynamics simulations.

Constructing a reliable damage criterion is a necessary and core task for the development of damage behavior models for soft materials. Damage criteria determine whether the damage happens and how the damage develops. Qi–Boyce model [97] and Gökşete–Miehe model model [100] adopted damage parameters to describe the relationship between the stress after damage occurs and the corresponding perfectly elastic stress; and the damage parameters depend on the SEDF and its historical maximum value. The model of Vernerey et al. [121] utilized a damage distribution function to describe the evolution of chain length distribution with damage. In the network alteration model of Markkannen et al. [26], based on a specific physical picture, the chain density decreases and the average chain length increases when the historical maximal deformation increases. All the damage models we mentioned are able to capture damage-induced stress softening behavior. Particularly, the models based on micro-sphere hyperelastic models, such as the model of Zhong et al. [27] and the Gökşete–Miehe model model [100], are able to further describe damage-induced residual deformation and anisotropy; in addition, Zhao model [28], Zhu and Zhong’s model [118] and model of Lavoie et al. [29] can describe the damage of multi-network elastomers.

The mechanics and physics of soft materials are immensely rich fields, and the current review, mostly focusing on hyperelasticity, visco-hyperelasticity, and damage, does not cover all the phenomena related to soft material behavior. We note that while the current state of constitutive models provides powerful means for modeling of conventional soft materials, the development of new synthetic and 3D-printed soft materials poses unique challenges and opportunities for the development of suitable constitutive models. Moreover, the constitutive models for soft materials are of essence for understanding the complex behavior of soft biological tissues and organs. Developing in vivo constitutive models for biological tissues and organs is of vital importance to develop reliable patient-specific models [122]. For instance, accurate modeling of the mechanical behaviors of the ventricular wall is helpful to develop a predictive patient-specific model for dysynchronous heart failure [123]. However, the difficulties to measure the accuracy in vivo properties pose a challenge to develop in vivo models. Interestingly, machine learning methods, especially deep learning, can be instrumental for the development of new constitutive models. One of the deep learning methods is called physics-informed neural network method, and it is started to be used to estimating unknown constitutive relations for given systems [124,125].

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant Nos. 11525210 and 91748209), and the Fundamental Research Funds for the Central Universities (Grant No. 2020XZZX005-02).

Appendix

In this Appendix, we will briefly introduce the deformation tensors used in this paper. Further comprehensive readings on this topic can be found from Ref. [37].

As shown in Fig. 18, the material point denoted by vector \( x \) in the reference configuration \( \Omega_0 \) is mapped by \( F \) into the position \( x \) in the deformed configuration \( \Omega \). Thus

\[
dx = F dx \tag{A1}\]

where

\[
F = \frac{\partial x}{\partial X} = \text{Grad} x \tag{A2}\]

is the definition of deformation gradient tensor.

The right and left Cauchy–Green deformation tensors are respectively defined by

\[
C = F^T F \tag{A3}\]

and

\[
B = FF^T \tag{A4}\]

The first invariant of the right (or left) Cauchy–Green deformation tensor which is frequently used in constitutive models is defined as

\[
I_1 = \text{tr}(C) = \text{tr}(B) \tag{A5}\]

![Fig. 18 Depiction of the reference configuration \( \Omega_0 \) mapping into the deformed configuration \( \Omega \)](image)


